

**COMPOSITIONS AND METHODS  
RELATING TO LUNG SPECIFIC GENES**

**INTRODUCTION**

This application claims the benefit of priority from  
5 U.S. Provisional Application Serial No. 60/228,378, filed  
August 28, 2000, which is herein incorporated in its  
entirety.

**FIELD OF THE INVENTION**

The present invention relates to newly identified  
10 nucleic acids and polypeptides present in normal and  
neoplastic lung cells, including fragments, variants and  
derivatives of the nucleic acids and polypeptides. The  
present invention also relates to antibodies to the  
polypeptides of the invention, as well as agonists and  
15 antagonists of the polypeptides of the invention. The  
invention also relates to compositions comprising the  
nucleic acids, polypeptides, antibodies, variants,  
derivatives, agonists and antagonists of the invention and  
methods for the use of these compositions. These uses  
20 include identifying, diagnosing, monitoring, staging,  
imaging and treating lung cancer and non-cancerous disease  
states in lung, identifying lung tissue, monitoring and  
modifying lung embryonic development and differentiation,  
and identifying and/or designing agonists and antagonists  
25 of polypeptides of the invention. The uses also include  
gene therapy, production of transgenic animals and cells,  
and production of engineered lung tissue for treatment and  
research.

## BACKGROUND OF THE INVENTION

Throughout the last hundred years, the incidence of lung cancer has steadily increased, so much so that now in many countries, it is the most common cancers. In fact, 5 lung cancer is the second most prevalent type of cancer for both men and women in the United States and is the most common cause of cancer death in both sexes. Lung cancer deaths have increased ten-fold in both men and women since 1930, primarily due to an increase in cigarette smoking, 10 but also due to an increased exposure to arsenic, asbestos, chromates, chloromethyl ethers, nickel, polycyclic aromatic hydrocarbons and other agents. See Scott, *Lung Cancer: A Guide to Diagnosis and Treatment*, Addicus Books (2000) and Alberg et al., in Kane et al. (eds.) *Biology of Lung* 15 *Cancer*, pp. 11-52, Marcel Dekker, Inc. (1998). Lung cancer may result from a primary tumor originating in the lung or a secondary tumor which has spread from another organ such as the bowel or breast. Although there are over a dozen types of lung cancer, over 90% fall into two categories: 20 small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). See Scott, *supra*. About 20-25% of all lung cancers are characterized as SCLC, while 70-80% are diagnosed as NSCLC. *Id.* A rare type of lung cancer is mesothelioma, which is generally caused by exposure to 25 asbestos, and which affects the pleura of the lung. Lung cancer is usually diagnosed or screened for by chest x-ray, CAT scans, PET scans, or by sputum cytology. A diagnosis of lung cancer is usually confirmed by biopsy of the tissue. *Id.*

30 SCLC tumors are highly metastatic and grow quickly. By the time a patient has been diagnosed with SCLC, the cancer has usually already spread to other parts of the body, including lymph nodes, adrenals, liver, bone, brain and bone marrow. See Scott, *supra*; Van Houtte et al. 35 (eds.), *Progress and Perspective in the Treatment of Lung*

Cancer, Springer-Verlag (1999). Because the disease has usually spread to such an extent that surgery is not an option, the current treatment of choice is chemotherapy plus chest irradiation. See Van Houtte, *supra*. The stage 5 of disease is a principal predictor of long-term survival. Less than 5% of patients with extensive disease that has spread beyond one lung and surrounding lymph nodes, live longer than two years. *Id.* However, the probability of five-year survival is three to four times higher if the 10 disease is diagnosed and treated when it is still in a limited stage, i.e., not having spread beyond one lung. *Id.*

NSCLC is generally divided into three types: squamous cell carcinoma, adenocarcinoma and large cell 15 carcinoma. Both squamous cell cancer and adenocarcinoma develop from the cells that line the airways; however, adenocarcinoma develops from the goblet cells that produce mucus. Large cell lung cancer has been thus named because the cells look large and rounded when viewed 20 microscopically, and generally are considered relatively undifferentiated. See Yesner, *Atlas of Lung Cancer*, Lippincott-Raven (1998).

Secondary lung cancer is a cancer initiated elsewhere in the body that has spread to the lungs. Cancers that 25 metastasize to the lung include, but are not limited to, breast cancer, melanoma, colon cancer and Hodgkin's lymphoma. Treatment for secondary lung cancer may depend upon the source of the original cancer. In other words, a lung cancer that originated from breast cancer may be more 30 responsive to breast cancer treatments and a lung cancer that originated from the colon cancer may be more responsive to colon cancer treatments.

The stage of a cancer indicates how far it has spread and is an important indicator of the prognosis. In 35 addition, staging is important because treatment is often

decided according to the stage of a cancer. SCLC is divided into two stages: limited disease, i.e., cancer that can only be seen in one lung and in nearby lymph nodes; and extensive disease, i.e., cancer that has spread 5 outside the lung to the chest or to other parts of the body. For most patients with SCLC, the disease has already progressed to lymph nodes or elsewhere in the body at the time of diagnosis. See Scott, *supra*. Even if spreading is not apparent on the scans, it is likely that some cancer 10 cells may have spread away and traveled through the bloodstream or lymph system. In general, chemotherapy with or without radiotherapy is often the preferred treatment. The initial scans and tests done at first will be used later to see how well a patient is responding to treatment.

15 In contrast, non-small cell cancer may be divided into four stages. Stage I is highly localized cancer with no cancer in the lymph nodes. Stage II cancer has spread to the lymph nodes at the top of the affected lung. Stage III cancer has spread near to where the cancer started. 20 This can be to the chest wall, the covering of the lung (pleura), the middle of the chest (mediastinum) or other lymph nodes. Stage IV cancer has spread to another part of the body. Stage I-III cancer is usually treated with surgery, with or without chemotherapy. Stage IV cancer is 25 usually treated with chemotherapy and/or palliative care.

A number of chromosomal and genetic abnormalities have been observed in lung cancer. In NSCLC, chromosomal aberrations have been described on 3p, 9p, 11p, 15p and 17p, and chromosomal deletions have been seen on 30 chromosomes 7, 11, 13 and 19. See Skarin (ed.), Multimodality Treatment of Lung Cancer, Marcel Dekker, Inc. (2000); Gemmill *et al.*, pp. 465-502, in Kane, *supra*; Bailey-Wilson *et al.*, pp. 53-98, in Kane, *supra*. Chromosomal abnormalities have been described on 1p, 3p, 35 5q, 6q, 8q, 13q and 17p in SCLC. *Id.* In addition, the

PAPERS REFERENCED

loss of the short arm of chromosome 3p has also been seen in greater than 90% of SCLC tumors and approximately 50% of NSCLC tumors. *Id.*

A number of oncogenes and tumor suppressor genes have 5 been implicated in lung cancer. See Mabry, pp. 391-412, in Kane, *supra* and Sclafani *et al.*, pp. 295-316, in Kane, *supra*. In both SCLC and NSCLC, the p53 tumor suppressor gene is mutated in over 50% of lung cancers. See Yesner, *supra*. Another tumor suppressor gene, FHIT, which is found 10 on chromosome 3p, is mutated by tobacco smoke. *Id.*; Skarin, *supra*. In addition, more than 95% of SCLCs and approximately 20-60% of NSCLCs have an absent or abnormal retinoblastoma (Rb) protein, another tumor suppressor gene. The *ras* oncogene (particularly *K-ras*) is mutated in 20-30% 15 of NSCLC specimens and the *c-erbB2* oncogene is expressed in 18% of stage 2 NSCLC and 60% of stage 4 NSCLC specimens. See Van Houtte, *supra*. Other tumor suppressor genes that are found in a region of chromosome 9, specifically in the region of 9p21, are deleted in many cancer cells, including 20 *p16<sup>INK4A</sup>* and *p15<sup>INK4B</sup>*. See Bailey-Wilson, *supra*; Sclafani *et al.*, *supra*. These tumor suppressor genes may also be implicated in lung cancer pathogenesis.

In addition, many lung cancer cells produce growth factors that may act in an autocrine fashion on lung cancer 25 cells. See Siegfried *et al.*, pp. 317-336, in Kane, *supra*; Moody, pp. 337-370, in Kane, *supra* and Heasley *et al.*, 371-390, in Kane, *supra*. In SCLC, many tumor cells produce gastrin-releasing peptide (GRP), which is a proliferative growth factor for these cells. See Skarin, *supra*. Many 30 NSCLC tumors express epidermal growth factor (EGF) receptors, allowing NSCLC cells to proliferate in response to EGF. Insulin-like growth factor (IGF-I) is elevated in greater than 95% of SCLC and greater than 80% of NSCLC tumors; it is thought to function as an autocrine growth

0250042204650

factor. *Id.* Finally, stem cell factor (SCF, also known as steel factor or kit ligand) and c-Kit (a proto-oncoprotein tyrosine kinase receptor for SCF) are both expressed at high levels in SCLC, and thus may form an autocrine loop 5 that increases proliferation. *Id.*

Although the majority of lung cancer cases are attributable to cigarette smoking, most smokers do not develop lung cancer. Epidemiological evidence has suggested that susceptibility to lung cancer may be 10 inherited in a Mendelian fashion, and thus have an inherited genetic component. Bailey-Wilson, *supra*. Thus, it is thought that certain allelic variants at some genetic loci may affect susceptibility to lung cancer. *Id.* One way to identify which allelic variants are likely to be 15 involved in lung cancer susceptibility, as well as susceptibility to other diseases, is to look at allelic variants of genes that are highly expressed in lung.

The lung is also susceptible to a number of other debilitating diseases, including, without limitation, 20 emphysema, pneumonia, cystic fibrosis and asthma. See Stockley (ed.), *Molecular Biology of the Lung*, Volume I: Emphysema and Infection, Birkhauser Verlag (1999), hereafter Stockley I, and Stockley (ed.), *Molecular Biology of the Lung*, Volume II: Asthma and Cancer, Birkhauser 25 Verlag (1999), hereafter Stockley II. The cause of many these disorders is still not well understood and there are few, if any, good treatment options for many of these noncancerous lung disorders. Thus, there remains a need to understand various noncancerous lung disorders and to 30 identify treatments for these diseases.

In yet another aspect, the development and differentiation of the lung tissue is important during embryonic development. All of the epithelial cells of the respiratory tract, including those of the lung and bronchi, 35 are derived from the primitive endodermal cells that line

059264 02304560

the embryonic outpouching. See Yesner, *supra*. During embryonic development, multipotent endodermal stem cells differentiate into many different types of specialized cells, which include ciliated cells for moving inhaled particles, goblet cells for producing mucus, Kulchitsky's cells for endocrine function, and Clara cells and type II pneumocytes for secreting surfactant protein. *Id.* Improper development and differentiation may cause respiratory disorders and distress in infants, particularly 10 in premature infants, whose lungs cannot produce sufficient surfactant when they are born. Further, some lung cancer cells, particularly small cell carcinomas, appear multipotent, and can spontaneously differentiate into a number of cell types, including small cell carcinoma, 15 adenocarcinoma and squamous cell carcinoma. *Id.* Thus, a better understanding of lung development and differentiation may help facilitate understanding of lung cancer initiation and progression.

Accordingly, there is a great need for more sensitive 20 and accurate methods for predicting whether a person is likely to develop lung cancer, for diagnosing lung cancer, for monitoring the progression of the disease, for staging the lung cancer, for determining whether the lung cancer has metastasized and for imaging the lung cancer. There is 25 also a need for better treatment of lung cancer. Further, there is also a great need for diagnosing and treating noncancerous lung disorders such as emphysema, pneumonia, lung infection, pulmonary fibrosis, cystic fibrosis and asthma. There is also a need for compositions and methods 30 of using them that can be used to identify lung tissue for forensic purposes and for determining whether a particular cell or tissue exhibits lung-specific characteristics.

In the present invention, methods are provided for detecting, diagnosing, monitoring, staging, 35 prognosticating, imaging and treating lung cancer via lung

specific genes referred to herein as LSGs. For purposes of the present invention, LSG refers, among other things, to native protein expressed by the gene comprising a polynucleotide sequence of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 5 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74, respectively. By "LSG" it is also meant herein polynucleotides which, due to degeneracy in genetic coding, comprise variations in nucleotide sequence as compared to SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 10 15, 16, 17, 18, 19, 20, or 74 but which still encode the same polypeptide. Exemplary amino acid sequences for LSG polypeptides are set forth in SEQ ID NO: 75, 76, 77, 78, 79, 80, 81, 82, 83 and 84. In the alternative, what is meant by LSG as used herein, means the native mRNA encoded 15 by the gene comprising the polynucleotide sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74, levels of the gene comprising the polynucleotide sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20 16, 17, 18, 19, 20 or 74 or levels of a polynucleotide which is capable of hybridizing under stringent conditions to the antisense sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 74.

Other objects, features, advantages and aspects of 25 the present invention will become apparent to those of skill in the art from the following description. It should be understood, however, that the following description and the specific examples, while indicating preferred embodiments of the invention are given by way of 30 illustration only. Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following description and from reading the other parts of the present disclosure.

## SUMMARY OF THE INVENTION

Toward these ends, and others, it is an object of the present invention to provide LSGs comprising a polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 5 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74, a protein expressed by a polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74, or a variant thereof which expresses the protein; or a polynucleotide which is capable of hybridizing under 10 stringent conditions to the antisense sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 74. Exemplary LSG polypeptides of the present invention are depicted in SEQ ID NO: 75, 76, 77, 78, 79, 80, 81, 82, 83 or 84.

15 It is another object of the present invention to provide a method for diagnosing the presence of lung cancer by analyzing for changes in levels of LSG in cells, tissues or bodily fluids compared with levels of LSG in preferably the same cells, tissues, or bodily fluid type of a normal 20 human control, wherein a change in levels of LSG in the patient versus the normal human control is associated with lung cancer.

Further provided is a method of diagnosing metastatic lung cancer in a patient having lung cancer which is not 25 known to have metastasized by identifying a human patient suspected of having lung cancer that has metastasized; analyzing a sample of cells, tissues, or bodily fluid from such patient for LSG; comparing the LSG levels in such cells, tissues, or bodily fluid with levels of LSG in 30 preferably the same cells, tissues, or bodily fluid type of a normal human control, wherein an increase in LSG levels in the patient versus the normal human control is associated with lung cancer which has metastasized.

Also provided by the invention is a method of staging 35 lung cancer in a human which has such cancer by identifying

a human patient having such cancer; analyzing a sample of cells, tissues, or bodily fluid from such patient for LSG; comparing LSG levels in such cells, tissues, or bodily fluid with levels of LSG in preferably the same cells,  
5 tissues, or bodily fluid type of a normal human control sample, wherein an increase in LSG levels in the patient versus the normal human control is associated with a cancer which is progressing and a decrease in the levels of LSG is associated with a cancer which is regressing or in  
10 remission.

Further provided is a method of monitoring lung cancer in a human having such cancer for the onset of metastasis. The method comprises identifying a human patient having such cancer that is not known to have  
15 metastasized; periodically analyzing a sample of cells, tissues, or bodily fluid from such patient for LSG; comparing the LSG levels in such cells, tissue, or bodily fluid with levels of LSG in preferably the same cells, tissues, or bodily fluid type of a normal human control  
20 sample, wherein an increase in LSG levels in the patient versus the normal human control is associated with a cancer which has metastasized.

Further provided is a method of monitoring the change in stage of lung cancer in a human having such cancer by  
25 looking at levels of LSG in a human having such cancer. The method comprises identifying a human patient having such cancer; periodically analyzing a sample of cells, tissues, or bodily fluid from such patient for LSG; comparing the LSG levels in such cells, tissue, or bodily  
30 fluid with levels of LSG in preferably the same cells, tissues, or bodily fluid type of a normal human control sample, wherein an increase in LSG levels in the patient versus the normal human control is associated with a cancer which is progressing and a decrease in the levels of LSG is

DRAFTING SHEET 1 OF 2

associated with a cancer which is regressing or in remission.

Further provided are methods of designing new therapeutic agents targeted to a LSG for use in imaging and 5 treating lung cancer. For example, in one embodiment, therapeutic agents such as antibodies targeted against LSG or fragments of such antibodies can be used to treat, detect or image localization of LSG in a patient for the purpose of detecting or diagnosing a disease or condition. 10 In this embodiment, an increase in the amount of labeled antibody detected as compared to normal tissue would be indicative of tumor metastases or growth. Such antibodies can be polyclonal, monoclonal, or omniclonal or prepared by molecular biology techniques. The term "antibody", as used 15 herein and throughout the instant specification is also meant to include aptamers and single-stranded oligonucleotides such as those derived from an *in vitro* evolution protocol referred to as SELEX and well known to those skilled in the art. Antibodies can be labeled with a 20 variety of detectable and therapeutic labels including, but not limited to, radioisotopes and paramagnetic metals. Therapeutic agents such as small molecules and antibodies which decrease the concentration and/or activity of LSG can also be used in the treatment of diseases characterized by 25 overexpression of LSG. Such agents can be readily identified in accordance with teachings herein.

Other objects, features, advantages and aspects of the present invention will become apparent to those of skill in the art from the following description. It should 30 be understood, however, that the following description and the specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only. Various changes and modifications within the spirit and scope of the disclosed invention will 35 become readily apparent to those skilled in the art from

reading the following description and from reading the other parts of the present disclosure.

#### GLOSSARY

The following illustrative explanations are provided  
5 to facilitate understanding of certain terms used  
frequently herein, particularly in the examples. The  
explanations are provided as a convenience and are not  
limitative of the invention.

ISOLATED means altered "by the hand of man" from its  
10 natural state; i.e., that, if it occurs in nature, it has  
been changed or removed from its original environment, or  
both.

For example, a naturally occurring polynucleotide or  
a polypeptide naturally present in a living animal in its  
15 natural state is not "isolated," but the same  
polynucleotide or polypeptide separated from the coexisting  
materials of its natural state is "isolated", as the term  
is employed herein. For example, with respect to  
polynucleotides, the term isolated means that it is  
20 separated from the chromosome and cell in which it  
naturally occurs.

As part of or following isolation, such  
polynucleotides can be joined to other polynucleotides,  
such as DNAs, for mutagenesis, to form fusion proteins, and  
25 for propagation or expression in a host, for instance. The  
isolated polynucleotides, alone or joined to other  
polynucleotides such as vectors, can be introduced into  
host cells, in culture or in whole organisms. When  
introduced into host cells in culture or in whole  
30 organisms, such DNAs still would be isolated, as the term  
is used herein, because they would not be in their  
naturally occurring form or environment. Similarly, the  
polynucleotides and polypeptides may occur in a  
composition, such as media formulations, solutions for

introduction of polynucleotides or polypeptides, for example, into cells, compositions or solutions for chemical or enzymatic reactions, for instance, which are not naturally occurring compositions, and, therein remain 5 isolated polynucleotides or polypeptides within the meaning of that term as it is employed herein.

OLIGONUCLEOTIDE(S) refers to relatively short polynucleotides. Often the term refers to single-stranded deoxyribonucleotides, but it can refer as well to single- or 10 double-stranded ribonucleotides, RNA:DNA hybrids and double-stranded DNAs, among others.

Oligonucleotides, such as single-stranded DNA probe oligonucleotides, often are synthesized by chemical methods, such as those implemented on automated 15 oligonucleotide synthesizers. However, oligonucleotides can be made by a variety of other methods, including *in vitro* recombinant DNA-mediated techniques and by expression of DNAs in cells and organisms.

Initially, chemically synthesized DNAs typically are 20 obtained without a 5' phosphate. The 5' ends of such oligonucleotides are not substrates for phosphodiester bond formation by ligation reactions that employ DNA ligases typically used to form recombinant DNA molecules. Where ligation of such oligonucleotides is desired, a phosphate 25 can be added by standard techniques, such as those that employ a kinase and ATP.

The 3' end of a chemically synthesized oligonucleotide generally has a free hydroxyl group and, in the presence of a ligase such as T4 DNA ligase, readily 30 will form a phosphodiester bond with a 5' phosphate of another polynucleotide, such as another oligonucleotide. As is well known, this reaction can be prevented selectively, where desired, by removing the 5' phosphates of the other polynucleotide(s) prior to ligation.

DRAFTING  
BY DEESE

POLYNUCLEOTIDE(S) generally refers to any polyribonucleotide or polydeoxribonucleotide and is inclusive of unmodified RNA or DNA as well as modified RNA or DNA. Thus, for instance, polynucleotides as used herein 5 refers to, among other things, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single- 10 stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, polynucleotide, as used herein, refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or 15 from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide.

20 As used herein, the term polynucleotide is also inclusive of DNAs or RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein.

25 Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein.

It will be appreciated that a great variety of  
30 modifications have been made to DNA and RNA that serve many  
useful purposes known to those of skill in the art. The  
term polynucleotide as it is employed herein embraces such  
chemically, enzymatically or metabolically modified forms  
of polynucleotides, as well as chemical forms of DNA and

RNA characteristic of viruses and cells, including simple and complex cells, *inter alia*.

POLYPEPTIDES, as used herein, includes all polypeptides as described below. The basic structure of 5 polypeptides is well known and has been described in innumerable textbooks and other publications in the art. In this context, the term is used herein to refer to any peptide or protein comprising two or more amino acids joined to each other in a linear chain by peptide bonds. 10 As used herein, the term refers to both short chains, which also commonly are referred to in the art as peptides, oligopeptides and oligomers, for example, and to longer chains, which generally are referred to in the art as proteins, of which there are many types. It will be 15 appreciated that polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally occurring amino acids, and that many amino acids, including the terminal amino acids, may be modified in a given polypeptide, either by natural processes such as 20 processing and other post-translational modifications, or by chemical modification techniques which are well known to the art. Even the common modifications that occur naturally in polypeptides are too numerous to list exhaustively here, but they are well described in basic 25 texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art.

Modifications which may be present in polypeptides of the present invention include, to name an illustrative few, 30 acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of a phosphotidylinositol, cross-linking, cyclization, disulfide 35

bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, 5 myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well known to those of skill 10 and have been described in great detail in the scientific literature. Several particularly common modifications including, but not limited to, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation are described 15 in most basic texts, such as, for instance PROTEINS STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as, for example, those provided by Wold, F., Posttranslational 20 Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York (1983); Seifter et al., Analysis for protein modifications and nonprotein cofactors, Meth. Enzymol. 182: 626-646 (1990) 25 and Rattan et al., Protein Synthesis: Posttranslational Modifications and Aging, Ann. N.Y. Acad. Sci. 663: 48-62 (1992).

It will be appreciated that the polypeptides of the present invention are not always entirely linear. Instead, 30 polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of posttranslation events including natural processing event and events brought about by human manipulation which do not occur naturally. Circular, 35 branched and branched circular polypeptides may be

synthesized by non-translation natural processes and by entirely synthetic methods, as well.

Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino and/or carboxyl group in a polypeptide by a covalent modification is common in naturally occurring and synthetic polypeptides and such modifications may be present in polypeptides of the present invention, as well. For instance, the amino terminal residue of polypeptides made in *E. coli*, prior to proteolytic processing, almost invariably will be N-formylmethionine.

The modifications that occur in a polypeptide often will be a function of how it is made. For polypeptides made by expressing a cloned gene in a host, for instance, the nature and extent of the modifications, in large part, will be determined by the host cell posttranslational modification capacity and the modification signals present in the polypeptide amino acid sequence. For instance, as is well known, glycosylation often does not occur in bacterial hosts such as *E. coli*. Accordingly, when glycosylation is desired, a polypeptide can be expressed in a glycosylating host, generally a eukaryotic cell. Insect cells often carry out the same posttranslational glycosylations as mammalian cells. Thus, insect cell expression systems have been developed to express efficiently mammalian proteins having native patterns of glycosylation, *inter alia*. Similar considerations apply to other modifications.

It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications.

In general, as used herein, the term polypeptide encompasses all such modifications, particularly those that

are present in polypeptides synthesized by expressing a polynucleotide in a host cell.

VARIANT(S) of polynucleotides or polypeptides, as the term is used herein, are polynucleotides or polypeptides 5 that differ from a reference polynucleotide or polypeptide, respectively.

With respect to variant polynucleotides, differences are generally limited so that the nucleotide sequences of the reference and the variant are closely similar overall 10 and, in many regions, identical. Thus, changes in the nucleotide sequence of the variant may be silent. That is, they may not alter the amino acids encoded by the polynucleotide. Where alterations are limited to silent changes of this type a variant will encode a polypeptide 15 with the same amino acid sequence as the reference.

Alternatively, changes in the nucleotide sequence of the variant may alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Such nucleotide changes may result in amino acid substitutions, additions, 20 deletions, fusions and truncations in the polypeptide encoded by the reference sequence.

With respect to variant polypeptides, differences are generally limited so that the sequences of the reference and the variant are closely similar overall and, in many 25 region, identical. For example, a variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions, fusions and truncations, which may be present in any combination.

RECEPTOR MOLECULE, as used herein, refers to 30 molecules which bind or interact specifically with LSG polypeptides of the present invention and is inclusive not only of classic receptors, which are preferred, but also other molecules that specifically bind to or interact with polypeptides of the invention (which also may be referred 35 to as "binding molecules" and "interaction molecules,"

CONFIDENTIAL

respectively and as "LSG binding or interaction molecules". Binding between polypeptides of the invention and such molecules, including receptor or binding or interaction molecules may be exclusive to polypeptides of the invention, which is very highly preferred, or it may be highly specific for polypeptides of the invention, which is highly preferred, or it may be highly specific to a group of proteins that includes polypeptides of the invention, which is preferred, or it may be specific to several groups of proteins at least one of which includes polypeptides of the invention.

Receptors also may be non-naturally occurring, such as antibodies and antibody-derived reagents that bind to polypeptides of the invention.

#### 15 DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to novel lung specific polypeptides and polynucleotides, referred to herein as LSGs, among other things, as described in greater detail below.

#### 20 *Polynucleotides*

In accordance with one aspect of the present invention, there are provided isolated LSG polynucleotides which encode LSG polypeptides.

Using the information provided herein, such as the 25 polynucleotide sequences set out in SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 74 a polynucleotide of the present invention encoding a LSG may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA from 30 cells of a human tumor as starting material.

Polynucleotides of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by

a combination thereof. The DNA may be double-stranded or single-stranded. Single-stranded DNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense  
5 strand.

The coding sequence which encodes the polypeptides may be identical to the coding sequence of the polynucleotides of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74. It also may  
10 be a polynucleotide with a different sequence, which, as a result of the redundancy (degeneracy) of the genetic code, encodes the same polypeptides as encoded by SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74.

15 Polynucleotides of the present invention, such as SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74 which encode these polypeptides may comprise the coding sequence for the mature polypeptide by itself. Polynucleotides of the present invention may  
20 also comprise the coding sequence for the mature polypeptide and additional coding sequences such as those encoding a leader or secretory sequence such as a pre-, or pro- or prepro-protein sequence. Polynucleotides of the present invention may also comprise the coding sequence of  
25 the mature polypeptide, with or without the aforementioned additional coding sequences, together with additional, non-coding sequences. Examples of additional non-coding sequences which may be incorporated into the polynucleotide of the present invention include, but are not limited to,  
30 introns and non-coding 5' and 3' sequences such as transcribed, non-translated sequences that play a role in transcription, mRNA processing including, for example, splicing and polyadenylation signals, ribosome binding and stability of mRNA, and additional coding sequence which  
35 codes for amino acids such as those which provide

additional functionalities. Thus, for instance, the polypeptide may be fused to a marker sequence such as a peptide which facilitates purification of the fused polypeptide. In certain preferred embodiments of this 5 aspect of the invention, the marker sequence is a hexa-histidine peptide, such as the tag provided in the pQE vector (Qiagen, Inc.), among others, many of which are commercially available. As described in Gentz et al. (Proc. Natl. Acad. Sci., USA 86: 821-824 (1989)), for instance, 10 hexa-histidine provides for convenient purification of the fusion protein. The HA tag corresponds to an epitope derived of influenza hemagglutinin protein (Wilson et al., Cell 37: 767 (1984)).

In accordance with the foregoing, the term 15 "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides which include a sequence encoding a polypeptide of the present invention, particularly SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74. Exemplary 20 polypeptides encoded by the polynucleotides are depicted in SEQ ID NO: 75, 76, 77, 78, 79, 80, 81, 82, 83 and 84. The term encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, interrupted by introns) together 25 with additional regions, that also may contain coding and/or non-coding sequences.

The present invention further relates to variants of the herein above described polynucleotides which encode for fragments, analogs and derivatives of the LSG polypeptides. 30 A variant of the polynucleotide may be a naturally occurring variant such as a naturally occurring allelic variant, or it may be a variant that is not known to occur naturally. Such non-naturally occurring variants of the polynucleotide may be made by mutagenesis techniques,

including those applied to polynucleotides, cells or organisms.

Among variants in this regard are variants that differ from the aforementioned polynucleotides by 5 nucleotide substitutions, deletions or additions. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding or non-coding regions or both. Alterations in the coding regions may produce conservative or non-conservative amino 10 acid substitutions, deletions or additions.

Among the particularly preferred embodiments of the invention in this regard are polynucleotides encoding polypeptides having the same amino acid sequence encoded by a LSG polynucleotide comprising SEQ ID NO: 1, 2, 3, 4, 5, 15 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74; variants, analogs, derivatives and fragments thereof, and fragments of the variants, analogs and derivatives. Exemplary polypeptides encoded by these polynucleotides are depicted in SEQ ID NO: 75, 76, 77, 78, 79, 80, 81, 82, 83 20 and 84. Further particularly preferred in this regard are LSG polynucleotides encoding polypeptide variants, analogs, derivatives and fragments, and variants, analogs and derivatives of the fragments, in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are 25 substituted, deleted or added, in any combination.

Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the LSG. Also especially preferred in this regard are conservative substitutions. Most highly 30 preferred are polynucleotides encoding polypeptides having the amino acid sequences as polypeptides encoded by SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74, without substitutions.

Further preferred embodiments of the invention are 35 LSG polynucleotides that are at least 70% identical to a

DEPARTMENT OF COMMERCE  
PATENT AND TRADEMARK OFFICE

polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74 and polynucleotides which are complementary to such polynucleotides. More preferred are LSG polynucleotides 5 that comprise a region that is at least 80% identical to a polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74. In this regard, LSG polynucleotides at least 90% identical to the same are particularly preferred, and among these 10 particularly preferred LSG polynucleotides, those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 15 99% being the most preferred.

Particularly preferred embodiments in this respect, moreover, are polynucleotides which encode polypeptides which retain substantially the same biological function or activity as the mature polypeptides encoded by a 20 polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74.

The present invention further relates to polynucleotides that hybridize to the herein above-described LSG sequences. In this regard, the present 25 invention especially relates to polynucleotides which hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% 30 identity between the sequences.

As discussed additionally herein regarding polynucleotide assays of the invention, for instance, polynucleotides of the invention as described herein, may be used as a hybridization probe for cDNA and genomic DNA 35 to isolate full-length cDNAs and genomic clones encoding

PENDING 02/28/2000  
DEX-0230

LSGs and to isolate cDNA and genomic clones of other genes that have a high sequence similarity to these LSGs. Such probes generally will comprise at least 15 bases. Preferably, such probes will have at least 30 bases and may 5 have at least 50 bases.

For example, the coding region of LSG of the present invention may be isolated by screening using an oligonucleotide probe synthesized from the known DNA sequence. A labeled oligonucleotide having a sequence 10 complementary to that of a gene of the present invention is used to screen a library of human cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes with.

The polynucleotides and polypeptides of the present  
15 invention may be employed as research reagents and  
materials for discovery of treatments and diagnostics to  
human disease, as further discussed herein relating to  
polynucleotide assays, *inter alia*.

The polynucleotides may encode a polypeptide which is  
20 the mature protein plus additional amino or carboxyl-  
terminal amino acids, or amino acids interior to the mature  
polypeptide (when the mature form has more than one  
polypeptide chain, for instance). Such sequences may play a  
role in processing of a protein from precursor to a mature  
25 form, may facilitate/protein trafficking, may prolong or  
shorten protein half-life or may facilitate manipulation of  
a protein for assay or production, among other things. As  
generally is the case *in situ*, the additional amino acids  
30 may be processed away from the mature protein by cellular  
enzymes.

A precursor protein having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed, such inactive precursors generally are activated.

activation. Generally, such precursors are called proproteins.

In sum, a polynucleotide of the present invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences which are not the leader sequences of a proprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

#### **Polypeptides**

The present invention further relates to LSG 15 polypeptides, preferably polypeptides encoded by a polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74. Exemplary polypeptides are depicted in SEQ ID NO: 75, 76, 77, 78, 79, 80, 81, 82, 83 or 84. The invention also relates to 20 fragments, analogs and derivatives of these polypeptides. The terms "fragment," "derivative" and "analog" when referring to the polypeptides of the present invention means a polypeptide which retains essentially the same biological function or activity as such polypeptides. 25 Thus, an analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature polypeptide.

The polypeptide of the present invention may be a recombinant polypeptide, a natural polypeptide or a 30 synthetic polypeptide. In certain preferred embodiments it is a recombinant polypeptide.

The fragment, derivative or analog of a polypeptide of or the present invention may be (I) one in which one or more of the amino acid residues are substituted with a 35 conserved or non-conserved amino acid residue (preferably a

02/22/2019 10:22:00 AM

conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code; (ii) one in which one or more of the amino acid residues includes a substituent group; (iii) one in which 5 the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol); or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a 10 sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

Among preferred variants are those that vary from a 15 reference by conservative amino acid substitutions. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among 20 the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic 25 residues Phe, Tyr.

The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The polypeptides of the present invention include the 30 polypeptides encoded by the polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74 (in particular the mature polypeptide) as well as polypeptides which have at least 75% similarity (preferably at least 75% identity), more preferably at 35 least 90% similarity (more preferably at least 90%

identity), still more preferably at least 95% similarity (still more preferably at least 95% identity), to a polypeptide encoded by SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74. Also 5 included are portions of such polypeptides generally containing at least 30 amino acids and more preferably at least 50 amino acids. Exemplary polypeptides are depicted in SEQ ID NO:75, 76, 77, 78, 79, 80, 81, 82, 83 or 84.

As known in the art "similarity" between two 10 polypeptides is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one polypeptide sequence with that of a second polypeptide.

Fragments or portions of the polypeptides of the present invention may be employed for producing the 15 corresponding full-length polypeptide by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length polypeptides. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of 20 the present invention.

#### ***Fragments***

Also among preferred embodiments of this aspect of the present invention are polypeptides comprising fragments of a polypeptide encoded by a polynucleotide of SEQ ID NO: 25 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74. In this regard a fragment is a polypeptide having an amino acid sequence that entirely is the same as part but not all of the amino acid sequence of the aforementioned LSG polypeptides and variants or 30 derivatives thereof.

Such fragments may be "free-standing," i.e., not part of or fused to other amino acids or polypeptides, or they may be contained within a larger polypeptide of which they form a part or region. When contained within a larger 35 polypeptide, the presently discussed fragments most

preferably form a single continuous region. However, several fragments may be comprised within a single larger polypeptide. For instance, certain preferred embodiments relate to a fragment of a LSG polypeptide of the present 5 comprised within a precursor polypeptide designed for expression in a host and having heterologous pre- and pro- polypeptide regions fused to the amino terminus of the LSG fragment and an additional region fused to the carboxyl terminus of the fragment. Therefore, fragments in one 10 aspect of the meaning intended herein, refers to the portion or portions of a fusion polypeptide or fusion protein derived from a LSG polypeptide.

As representative examples of polypeptide fragments of the invention, there may be mentioned those which have 15 from about 15 to about 139 amino acids. In this context "about" includes the particularly recited range and ranges larger or smaller by several, a few, 5, 4, 3, 2 or 1 amino acid at either extreme or at both extremes. Highly preferred in this regard are the recited ranges plus or 20 minus as many as 5 amino acids at either or at both extremes. Particularly highly preferred are the recited ranges plus or minus as many as 3 amino acids at either or at both the recited extremes. Especially preferred are ranges plus or minus 1 amino acid at either or at both 25 extremes or the recited ranges with no additions or deletions. Most highly preferred of all in this regard are fragments from about 15 to about 45 amino acids.

Among especially preferred fragments of the invention are truncation mutants of the LSG polypeptides. Truncation 30 mutants include LSG polypeptides having an amino acid sequence encoded by a polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74 or variants or derivatives thereof, except for 35 deletion of a continuous series of residues (that is, a continuous region, part or portion) that includes the amino

terminus, or a continuous series of residues that includes the carboxyl terminus or, as in double truncation mutants, deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl 5 terminus. Fragments having the size ranges set out herein also are preferred embodiments of truncation fragments, which are especially preferred among fragments generally.

Also preferred in this aspect of the invention are fragments characterized by structural or functional 10 attributes of the LSG polypeptides of the present invention. Preferred embodiments of the invention in this regard include fragments that comprise alpha-helix and alpha-helix forming regions ("alpha-regions"), beta-sheet and beta-sheet-forming regions ("beta-regions"), turn and 15 turn-forming regions ("turn-regions"), coil and coil-forming regions ("coil-regions"), hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions and high antigenic index regions of the LSG 20 polypeptides of the present invention. Regions of the aforementioned types are identified routinely by analysis of the amino acid sequences encoded by the polynucleotides of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74. Preferred regions 25 include Garnier-Robson alpha-regions, beta-regions, turn-regions and coil-regions, Chou-Fasman alpha-regions, beta-regions and turn-regions, Kyte-Doolittle hydrophilic regions and hydrophilic regions, Eisenberg alpha and beta amphipathic regions, Karplus-Schulz flexible regions, Emini 30 surface-forming regions and Jameson-Wolf high antigenic index regions. Among highly preferred fragments in this regard are those that comprise regions of LSGs that combine several structural features, such as several of the features set out above. In this regard, the regions 35 defined by selected residues of a LSG polypeptide which all

02/28/2019 10:50:50

are characterized by amino acid compositions highly characteristic of turn-regions, hydrophilic regions, flexible-regions, surface-forming regions, and high antigenic index-regions, are especially highly preferred regions. Such regions may be comprised within a larger polypeptide or may be by themselves a preferred fragment of the present invention, as discussed above. It will be appreciated that the term "about" as used in this paragraph has the meaning set out above regarding fragments in general.

Further preferred regions are those that mediate activities of LSG polypeptides. Most highly preferred in this regard are fragments that have a chemical, biological or other activity of a LSG polypeptide, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Highly preferred in this regard are fragments that contain regions that are homologs in sequence, or in position, or in both sequence and to active regions of related polypeptides, and which include lung specific-binding proteins. Among particularly preferred fragments in these regards are truncation mutants, as discussed above.

It will be appreciated that the invention also relates to polynucleotides encoding the aforementioned fragments, polynucleotides that hybridize to polynucleotides encoding the fragments, particularly those that hybridize under stringent conditions, and polynucleotides such as PCR primers for amplifying polynucleotides that encode the fragments. In these regards, preferred polynucleotides are those that correspond to the preferred fragments, as discussed above.

#### ***Fusion Proteins***

In one embodiment of the present invention, the LSG polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a

variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See also EP A 394,827; Traunecker, et al., Nature 331: 84-86 5 (1988)) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time *in vivo*. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can 10 increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of these types of 15 fusion proteins described above can be made in accordance with well known protocols.

For example, a LSG polypeptide can be fused to an IgG molecule via the following protocol. Briefly, the human Fc portion of the IgG molecule is PCR amplified using primers 20 that span the 5' and 3' ends of the sequence. These primers also have convenient restriction enzyme sites that facilitate cloning into an expression vector, preferably a mammalian expression vector. For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be 25 ligated into the BamHI cloning site. In this protocol, the 3' BamHI site must be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI thereby linearizing the vector, and a LSG polynucleotide of the present invention is ligated into this BamHI site. It 30 is preferred that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second 35 signal peptide. Alternatively, if the naturally occurring

signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e. g., WO 96/34891.)

***Diagnostic Assays***

5 The present invention also relates to diagnostic assays and methods, both quantitative and qualitative for detecting, diagnosing, monitoring, staging and prognosticating cancers by comparing levels of LSG in a human patient with those of LSG in a normal human control.

10 For purposes of the present invention, what is meant by LSG levels is, among other things, native protein expressed by a gene comprising the polynucleotide sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74. Exemplary polypeptides encoded by these

15 polynucleotides are depicted in SEQ ID NO: 75, 76, 77, 78, 79, 80, 81, 82, 83 and 84. By "LSG" it is also meant herein polynucleotides which, due to degeneracy in genetic coding, comprise variations in nucleotide sequence as compared to SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11,

20 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74 but which still encode the same protein. The native protein being detected may be whole, a breakdown product, a complex of molecules or chemically modified. In the alternative, what is meant by LSG as used herein, means the native mRNA encoded by a

25 polynucleotide sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74, levels of the gene comprising the polynucleotide sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74, or levels of a

30 polynucleotide which is capable of hybridizing under stringent conditions to the antisense sequence of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74. Such levels are preferably determined in at least one of cells, tissues and/or bodily

35 fluids, including determination of normal and abnormal

levels. Thus, for instance, a diagnostic assay in accordance with the invention for diagnosing overexpression of LSG protein compared to normal control bodily fluids, cells, or tissue samples may be used to diagnose the 5 presence of lung cancer.

All the methods of the present invention may optionally include determining the levels of other cancer markers as well as LSG. Other cancer markers, in addition to LSG, useful in the present invention will depend on the 10 cancer being tested and are known to those of skill in the art.

The present invention provides methods for diagnosing the presence of lung cancer by analyzing for changes in levels of LSG in cells, tissues or bodily fluids compared 15 with levels of LSG in cells, tissues or bodily fluids of preferably the same type from a normal human control, wherein an increase in levels of LSG in the patient versus the normal human control is associated with the presence of lung cancer.

20 Without limiting the instant invention, typically, for a quantitative diagnostic assay a positive result indicating the patient being tested has cancer is one in which cells, tissues or bodily fluid levels of the cancer marker, such as LSG, are at least two times higher, and 25 most preferably are at least five times higher, than in preferably the same cells, tissues or bodily fluid of a normal human control.

The present invention also provides a method of diagnosing metastatic lung cancer in a patient having lung 30 cancer which has not yet metastasized for the onset of metastasis. In the method of the present invention, a human cancer patient suspected of having lung cancer which may have metastasized (but which was not previously known to have metastasized) is identified. This is accomplished 35 by a variety of means known to those of skill in the art.

In the present invention, determining the presence of LSG levels in cells, tissues or bodily fluid, is particularly useful for discriminating between lung cancer which has not metastasized and lung cancer which has metastasized. Existing techniques have difficulty discriminating between lung cancer which has metastasized and lung cancer which has not metastasized and proper treatment selection is often dependent upon such knowledge.

In the present invention, the cancer marker levels measured in such cells, tissues or bodily fluid is LSG, and are compared with levels of LSG in preferably the same cells, tissue or bodily fluid type of a normal human control. That is, if the cancer marker being observed is just LSG in serum, this level is preferably compared with the level of LSG in serum of a normal human control. An increase in the LSG in the patient versus the normal human control is associated with lung cancer which has metastasized.

Without limiting the instant invention, typically, for a quantitative diagnostic assay a positive result indicating the cancer in the patient being tested or monitored has metastasized is one in which cells, tissues or bodily fluid levels of the cancer marker, such as LSG, are at least two times higher, and most preferably are at least five times higher, than in preferably the same cells, tissues or bodily fluid of a normal patient.

Normal human control as used herein includes a human patient without cancer and/or non cancerous samples from the patient; in the methods for diagnosing or monitoring for metastasis, normal human control may preferably also include samples from a human patient that is determined by reliable methods to have lung cancer which has not metastasized.

*Staging*

The invention also provides a method of staging lung cancer in a human patient. The method comprises identifying a human patient having such cancer and 5 analyzing cells, tissues or bodily fluid from such human patient for LSG. The LSG levels determined in the patient are then compared with levels of LSG in preferably the same cells, tissues or bodily fluid type of a normal human control, wherein an increase in LSG levels in the human 10 patient versus the normal human control is associated with a cancer which is progressing and a decrease in the levels of LSG (but still increased over true normal levels) is associated with a cancer which is regressing or in remission.

**15 Monitoring**

Further provided is a method of monitoring lung cancer in a human patient having such cancer for the onset of metastasis. The method comprises identifying a human patient having such cancer that is not known to have 20 metastasized; periodically analyzing cells, tissues or bodily fluid from such human patient for LSG; and comparing the LSG levels determined in the human patient with levels of LSG in preferably the same cells, tissues or bodily fluid type of a normal human control, wherein an increase 25 in LSG levels in the human patient versus the normal human control is associated with a cancer which has metastasized. In this method, normal human control samples may also include prior patient samples.

Further provided by this invention is a method of 30 monitoring the change in stage of lung cancer in a human patient having such cancer. The method comprises identifying a human patient having such cancer; periodically analyzing cells, tissues or bodily fluid from such human patient for LSG; and comparing the LSG levels 35 determined in the human patient with levels of LSG in

02/20/2014 10:56:50

preferably the same cells, tissues or bodily fluid type of a normal human control, wherein an increase in LSG levels in the human patient versus the normal human control is associated with a cancer which is progressing in stage and 5 a decrease in the levels of LSG is associated with a cancer which is regressing in stage or in remission. In this method, normal human control samples may also include prior patient samples.

Monitoring a patient for onset of metastasis is 10 periodic and preferably done on a quarterly basis. However, this may be done more or less frequently depending on the cancer, the particular patient, and the stage of the cancer.

***Prognostic Testing and Clinical Trial Monitoring***

15 The methods described herein can further be utilized as prognostic assays to identify subjects having or at risk of developing a disease or disorder associated with increased levels of LSG. The present invention provides a method in which a test sample is obtained from a human 20 patient and LSG is detected. The presence of higher LSG levels as compared to normal human controls is diagnostic for the human patient being at risk for developing cancer, particularly lung cancer.

The effectiveness of therapeutic agents to decrease 25 expression or activity of the LSGs of the invention can also be monitored by analyzing levels of expression of the LSGs in a human patient in clinical trials or in *in vitro* screening assays such as in human cells. In this way, the gene expression pattern can serve as a marker, indicative 30 of the physiological response of the human patient, or cells as the case may be, to the agent being tested.

***Detection of genetic lesions or mutations***

The methods of the present invention can also be used to detect genetic lesions or mutations in LSG, thereby 35 determining if a human with the genetic lesion is at risk

for lung cancer or has lung cancer. Genetic lesions can be detected, for example, by ascertaining the existence of a deletion and/or addition and/or substitution of one or more nucleotides from the LSGs of this invention, a chromosomal rearrangement of LSG, aberrant modification of LSG (such as of the methylation pattern of the genomic DNA), the presence of a non-wild type splicing pattern of a mRNA transcript of LSG, allelic loss of LSG, and/or inappropriate post-translational modification of LSG protein. Methods to detect such lesions in the LSG of this invention are known to those of skill in the art.

For example, in one embodiment, alterations in a gene corresponding to a LSG polynucleotide of the present invention are determined via isolation of RNA from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. See, e.g. Sambrook et al. (MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)), which is illustrative of the many laboratory manuals that detail these techniques. The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or 74. PCR conditions typically consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252: 706 (1991). PCR products are sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase (Epicentre Technologies). The intron-exon borders of selected exons are also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations are then cloned and sequenced to validate the results of the

direct sequencing. PCR products are cloned into T-tailed vectors as described in Holton, T. A. and Graham, M. W., Nucleic Acids Research, 19 : 1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected 5 individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements can also be observed as a method of determining alterations in a gene corresponding to a polynucleotide. In this method, genomic clones are 10 nick-translated with digoxigenin deoxy-uridine 5' triphosphate (Boehringer Manheim), and FISH is performed as described in Johnson, C. et al., Methods Cell Biol. 35: 73-99 (1991). Hybridization with a labeled probe is carried out using a vast excess of human DNA for specific 15 hybridization to the corresponding genomic locus. Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C-and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, 20 Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters (Johnson et al., Genet. Anal. Tech. Appl., 8: 75 (1991)). Image collection, analysis and chromosomal fractional length measurements are 25 performed using the ISee Graphical Program System (Inovision Corporation, Durham, NC). Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic 30 marker for an associated disease.

#### **Assay Techniques**

Assay techniques that can be used to determine levels of gene expression (including protein levels), such as LSG of the present invention, in a sample derived from a 35 patient are well known to those of skill in the art. Such

assay methods include, without limitation, radioimmunoassays, reverse transcriptase PCR (RT-PCR) assays, immunohistochemistry assays, *in situ* hybridization assays, competitive-binding assays, Western Blot analyses,

5 ELISA assays and proteomic approaches: two-dimensional gel electrophoresis (2D electrophoresis) and non-gel based approaches such as mass spectrometry or protein interaction profiling. Among these, ELISAs are frequently preferred to diagnose a gene's expressed protein in biological fluids.

10 An ELISA assay initially comprises preparing an antibody, if not readily available from a commercial source, specific to LSG, preferably a monoclonal antibody. In addition a reporter antibody generally is prepared which binds specifically to LSG. The reporter antibody is  
15 attached to a detectable reagent such as radioactive, fluorescent or enzymatic reagent, for example horseradish peroxidase enzyme or alkaline phosphatase.

To carry out the ELISA, antibody specific to LSG is incubated on a solid support, e.g. a polystyrene dish, that  
20 binds the antibody. Any free protein binding sites on the dish are then covered by incubating with a non-specific protein such as bovine serum albumin. Next, the sample to be analyzed is incubated in the dish, during which time LSG binds to the specific antibody attached to the polystyrene  
25 dish. Unbound sample is washed out with buffer. A reporter antibody specifically directed to LSG and linked to a detectable reagent such as horseradish peroxidase is placed in the dish resulting in binding of the reporter antibody to any monoclonal antibody bound to LSG. Unattached  
30 reporter antibody is then washed out. Reagents for peroxidase activity, including a colorimetric substrate are then added to the dish. Immobilized peroxidase, linked to LSG antibodies, produces a colored reaction product. The amount of color developed in a given time period is  
35 proportional to the amount of LSG protein present in the

P0230-20240500

sample. Quantitative results typically are obtained by reference to a standard curve.

A competition assay can also be employed wherein antibodies specific to LSG are attached to a solid support 5 and labeled LSG and a sample derived from the host are passed over the solid support. The amount of label detected which is attached to the solid support can be correlated to a quantity of LSG in the sample.

Using all or a portion of a nucleic acid sequence of 10 LSG of the present invention as a hybridization probe, nucleic acid methods can also be used to detect LSG mRNA as a marker for lung cancer. Polymerase chain reaction (PCR) and other nucleic acid methods, such as ligase chain reaction (LCR) and nucleic acid sequence based 15 amplification (NASBA), can be used to detect malignant cells for diagnosis and monitoring of various malignancies. For example, reverse-transcriptase PCR (RT-PCR) is a powerful technique which can be used to detect the presence of a specific mRNA population in a complex mixture of 20 thousands of other mRNA species. In RT-PCR, an mRNA species is first reverse transcribed to complementary DNA (cDNA) with use of the enzyme reverse transcriptase; the cDNA is then amplified as in a standard PCR reaction. RT-PCR can thus reveal by amplification the presence of a 25 single species of mRNA. Accordingly, if the mRNA is highly specific for the cell that produces it, RT-PCR can be used to identify the presence of a specific type of cell.

Hybridization to clones or oligonucleotides arrayed on a solid support (i.e. gridding) can be used to both 30 detect the expression of and quantitate the level of expression of that gene. In this approach, a cDNA encoding the LSG gene is fixed to a substrate. The substrate may be of any suitable type including but not limited to glass, nitrocellulose, nylon or plastic. At least a portion of 35 the DNA encoding the LSG gene is attached to the substrate

and then incubated with the analyte, which may be RNA or a complementary DNA (cDNA) copy of the RNA, isolated from the tissue of interest. Hybridization between the substrate bound DNA and the analyte can be detected and quantitated 5 by several means including but not limited to radioactive labeling or fluorescence labeling of the analyte or a secondary molecule designed to detect the hybrid.

Quantitation of the level of gene expression can be done by comparison of the intensity of the signal from the analyte 10 compared with that determined from known standards. The standards can be obtained by *in vitro* transcription of the target gene, quantitating the yield, and then using that material to generate a standard curve.

Of the proteomic approaches, 2D electrophoresis is a 15 technique well known to those in the art. Isolation of individual proteins from a sample such as serum is accomplished using sequential separation of proteins by different characteristics usually on polyacrylamide gels. First, proteins are separated by size using an electric 20 current. The current acts uniformly on all proteins, so smaller proteins move farther on the gel than larger proteins. The second dimension applies a current perpendicular to the first and separates proteins not on the basis of size but on the specific electric charge 25 carried by each protein. Since no two proteins with different sequences are identical on the basis of both size and charge, the result of a 2D separation is a square gel in which each protein occupies a unique spot. Analysis of the spots with chemical or antibody probes, or subsequent 30 protein microsequencing can reveal the relative abundance of a given protein and the identity of the proteins in the sample.

The above tests can be carried out on samples derived from a variety of cells, bodily fluids and/or tissue 35 extracts such as homogenates or solubilized tissue obtained

from a patient. Tissue extracts are obtained routinely from tissue biopsy and autopsy material. Bodily fluids useful in the present invention include blood, urine, saliva or any other bodily secretion or derivative thereof.

5 By blood it is meant to include whole blood, plasma, serum or any derivative of blood.

***In Vivo Targeting of LSG/Lung Cancer Therapy***

Identification of this LSG is also useful in the rational design of new therapeutics for imaging and

10 treating cancers, and in particular lung cancer. For example, in one embodiment, antibodies which specifically bind to LSG can be raised and used *in vivo* in patients suspected of suffering from lung cancer. Antibodies which specifically bind LSG can be injected into a patient

15 suspected of having lung cancer for diagnostic and/or therapeutic purposes. Thus, another aspect of the present invention provides for a method for preventing the onset and treatment of lung cancer in a human patient in need of such treatment by administering to the patient an effective 20 amount of antibody. By "effective amount" it is meant the amount or concentration of antibody needed to bind to the target antigens expressed on the tumor to cause tumor shrinkage for surgical removal, or disappearance of the tumor. The binding of the antibody to the overexpressed

25 LSG is believed to cause the death of the cancer cell expressing such LSG. The preparation and use of antibodies for *in vivo* diagnosis and treatment is well known in the art. For example, antibody-chelators labeled with Indium-111 have been described for use in the

30 radioimmunoscintographic imaging of carcinoembryonic antigen expressing tumors (Sumerdon et al. Nucl. Med. Biol. 1990 17:247-254). In particular, these antibody-chelators have been used in detecting tumors in patients suspected of having recurrent colorectal cancer (Griffin et al. J. Clin. 35 Onc. 1991 9:631-640). Antibodies with paramagnetic ions as

100-200-300-400-500-600-700-800-900

labels for use in magnetic resonance imaging have also been described (Lauffer, R.B. Magnetic Resonance in Medicine 1991 22:339-342). Antibodies directed against LSG can be used in a similar manner. Labeled antibodies which 5 specifically bind LSG can be injected into patients suspected of having lung cancer for the purpose of diagnosing or staging of the disease status of the patient. The label used will be selected in accordance with the imaging modality to be used. For example, radioactive 10 labels such as Indium-111, Technetium-99m or Iodine-131 can be used for planar scans or single photon emission computed tomography (SPECT). Positron emitting labels such as Fluorine-19 can be used in positron emission tomography. Paramagnetic ions such as Gadlinium (III) or Manganese (II) 15 can be used in magnetic resonance imaging (MRI). Presence of the label, as compared to imaging of normal tissue, permits determination of the spread of the cancer. The amount of label within an organ or tissue also allows determination of the presence or absence of cancer in that 20 organ or tissue.

Antibodies which can be used in *in vivo* methods include polyclonal, monoclonal and omniclonal antibodies and antibodies prepared via molecular biology techniques. Antibody fragments and aptamers and single-stranded 25 oligonucleotides such as those derived from an *in vitro* evolution protocol referred to as SELEX and well known to those skilled in the art can also be used.

#### **Screening Assays**

The present invention also provides methods for 30 identifying modulators which bind to LSG protein or have a modulatory effect on the expression or activity of LSG protein. Modulators which decrease the expression or activity of LSG protein are believed to be useful in treating lung cancer. Such screening assays are known to

those of skill in the art and include, without limitation, cell-based assays and cell free assays.

Small molecules predicted via computer imaging to specifically bind to regions of LSG can also be designed, 5 synthesized and tested for use in the imaging and treatment of lung cancer. Further, libraries of molecules can be screened for potential anticancer agents by assessing the ability of the molecule to bind to the LSGs identified herein. Molecules identified in the library as being 10 capable of binding to LSG are key candidates for further evaluation for use in the treatment of lung cancer. In a preferred embodiment, these molecules will downregulate expression and/or activity of LSG in cells.

***Adoptive Immunotherapy and Vaccines***

15 Adoptive immunotherapy of cancer refers to a therapeutic approach in which immune cells with an antitumor reactivity are administered to a tumor-bearing host, with the aim that the cells mediate either directly or indirectly, the regression of an established tumor. 20 Transfusion of lymphocytes, particularly T lymphocytes, falls into this category and investigators at the National Cancer Institute (NCI) have used autologous reinfusion of peripheral blood lymphocytes or tumor-infiltrating lymphocytes (TIL), T cell cultures from biopsies of 25 subcutaneous lymph nodules, to treat several human cancers (Rosenberg, S. A., U.S. Patent No. 4,690,914, issued Sep. 1, 1987; Rosenberg, S. A., et al., 1988, N. England J. Med. 319:1676-1680).

The present invention relates to compositions and 30 methods of adoptive immunotherapy for the prevention and/or treatment of primary and metastatic lung cancer in humans using macrophages sensitized to the antigenic LSG molecules, with or without non-covalent complexes of heat shock protein (hsp). Antigenicity or immunogenicity of the 35 LSG is readily confirmed by the ability of the LSG protein

or a fragment thereof to raise antibodies or educate naive effector cells, which in turn lyse target cells expressing the antigen (or epitope).

Cancer cells are, by definition, abnormal and contain 5 proteins which should be recognized by the immune system as foreign since they are not present in normal tissues. However, the immune system often seems to ignore this abnormality and fails to attack tumors. The foreign LSG proteins that are produced by the cancer cells can be used 10 to reveal their presence. The LSG is broken into short fragments, called tumor antigens, which are displayed on the surface of the cell. These tumor antigens are held or presented on the cell surface by molecules called MHC, of which there are two types: class I and II. Tumor antigens 15 in association with MHC class I molecules are recognized by cytotoxic T cells while antigen-MHC class II complexes are recognized by a second subset of T cells called helper cells. These cells secrete cytokines which slow or stop tumor growth and help another type of white blood cell, B 20 cells, to make antibodies against the tumor cells.

In adoptive immunotherapy, T cells or other antigen presenting cells (APCs) are stimulated outside the body (*ex vivo*), using the tumor specific LSG antigen. The stimulated cells are then reinfused into the patient where 25 they attack the cancerous cells. Research has shown that using both cytotoxic and helper T cells is far more effective than using either subset alone. Additionally, the LSG antigen may be complexed with heat shock proteins to stimulate the APCs as described in U.S. Patent No.

30 5,985,270.

The APCs can be selected from among those antigen presenting cells known in the art including, but not limited to, macrophages, dendritic cells, B lymphocytes, and a combination thereof, and are preferably macrophages. 35 In a preferred use, wherein cells are autologous to the

individual, autologous immune cells such as lymphocytes, macrophages or other APCs are used to circumvent the issue of whom to select as the donor of the immune cells for adoptive transfer. Another problem circumvented by use of 5 autologous immune cells is graft versus host disease which can be fatal if unsuccessfully treated.

In adoptive immunotherapy with gene therapy, DNA of the LSG can be introduced into effector cells similarly as in conventional gene therapy. This can enhance the 10 cytotoxicity of the effector cells to tumor cells as they have been manipulated to produce the antigenic protein resulting in improvement of the adoptive immunotherapy.

LSG antigens of this invention are also useful as components of lung cancer vaccines. The vaccine comprises 15 an immunogenically stimulatory amount of a LSG antigen. Immunogenically stimulatory amount refers to that amount of antigen that is able to invoke the desired immune response in the recipient for the amelioration, or treatment of lung cancer. Effective amounts may be determined empirically by 20 standard procedures well known to those skilled in the art.

The LSG antigen may be provided in any one of a number of vaccine formulations which are designed to induce the desired type of immune response, e.g., antibody and/or cell mediated. Such formulations are known in the art and 25 include, but are not limited to, formulations such as those described in U.S. Patent 5,585,103. Vaccine formulations of the present invention used to stimulate immune responses can also include pharmaceutically acceptable adjuvants.

***Vectors, host cells, expression***

30 The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

Host cells can be genetically engineered to incorporate LSG polynucleotides and express LSG polypeptides of the present invention. For instance, LSG polynucleotides may be introduced into host cells using 5 well known techniques of infection, transduction, transfection, transvection and transformation. The LSG polynucleotides may be introduced alone or with other polynucleotides. Such other polynucleotides may be introduced independently, co-introduced or introduced 10 joined to the LSG polynucleotides of the invention.

For example, LSG polynucleotides of the invention may be transfected into host cells with another, separate, polynucleotide encoding a selectable marker, using standard techniques for co-transfection and selection in, for 15 instance, mammalian cells. In this case, the polynucleotides generally will be stably incorporated into the host cell genome.

Alternatively, the LSG polynucleotide may be joined to a vector containing a selectable marker for propagation 20 in a host. The vector construct may be introduced into host cells by the aforementioned techniques. Generally, a plasmid vector is introduced as DNA in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. Electroporation also may be used to 25 introduce LSG polynucleotides into a host. If the vector is a virus, it may be packaged *in vitro* or introduced into a packaging cell and the packaged virus may be transduced into cells. A wide variety of well known techniques conducted routinely by those of skill in the art are 30 suitable for making LSG polynucleotides and for introducing LSG polynucleotides into cells in accordance with this aspect of the invention. Such techniques are reviewed at length in reference texts such as Sambrook et al., previously cited herein.

Vectors which may be used in the present invention include, for example, plasmid vectors, single- or double-stranded phage vectors, and single- or double-stranded RNA or DNA viral vectors. Such vectors may be introduced into 5 cells as polynucleotides, preferably DNA, by well known techniques for introducing DNA and RNA into cells. The vectors, in the case of phage and viral vectors, also may be and preferably are introduced into cells as packaged or encapsidated virus by well known techniques for infection 10 and transduction. Viral vectors may be replication competent or replication defective. In the latter case viral propagation generally will occur only in complementing host cells.

Preferred vectors for expression of polynucleotides and polypeptides of the present invention include, but are not limited to, vectors comprising cis-acting control regions effective for expression in a host operatively linked to the polynucleotide to be expressed. Appropriate trans-acting factors either are supplied by the host, supplied by a complementing vector or supplied by the vector itself upon introduction into the host.

In certain preferred embodiments in this regard, the vectors provide for specific expression. Such specific expression may be inducible expression or expression only 25 in certain types of cells or both inducible and cell-specific. Particularly preferred among inducible vectors are vectors that can be induced to express by environmental factors that are easy to manipulate, such as temperature and nutrient additives. A variety of vectors suitable to 30 this aspect of the invention, including constitutive and inducible expression vectors for use in prokaryotic and eukaryotic hosts, are well known and employed routinely by those of skill in the art.

The engineered host cells can be cultured in conventional nutrient media which may be modified as

appropriate for, *inter alia*, activating promoters, selecting transformants or amplifying genes. Culture conditions such as temperature, pH and the like, previously used with the host cell selected for expression, generally 5 will be suitable for expression of LSG polypeptides of the present invention.

A great variety of expression vectors can be used to express LSG polypeptides of the invention. Such vectors include chromosomal, episomal and virus-derived vectors.

10 Vectors may be derived from bacterial plasmids, from bacteriophage, from yeast episomes, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and 15 from combinations thereof such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. All may be used for expression in accordance with this aspect of the present invention. Generally, any vector suitable to maintain, propagate or express 20 polynucleotides to express a polypeptide in a host may be used for expression in this regard.

The appropriate DNA sequence may be inserted into the vector by any of a variety of well-known and routine techniques. In general, a DNA sequence for expression is 25 joined to an expression vector by cleaving the DNA sequence and the expression vector with one or more restriction endonucleases and then joining the restriction fragments together using T4 DNA ligase. Procedures for restriction and ligation that can be used to this end are well known 30 and routine to those of skill. Suitable procedures in this regard, and for constructing expression vectors using alternative techniques, which also are well known and routine to those skill, are set forth in great detail in Sambrook et al. cited elsewhere herein.

1025220-A20070609

The DNA sequence in the expression vector is operatively linked to appropriate expression control sequence(s), including, for instance, a promoter to direct mRNA transcription. Representative promoters include the 5 phage lambda PL promoter, the *E. coli* lac, trp and tac promoters, the SV40 early and late promoters, and promoters of retroviral LTRs, to name just a few of the well-known promoters. It will be understood that numerous promoters not mentioned are also suitable for use in this aspect of 10 the invention and are well known and readily may be employed by those of skill in the manner illustrated by the discussion and the examples herein.

In general, expression constructs will contain sites for transcription initiation and termination, and, in the 15 transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will include a translation initiating AUG at the beginning and a termination codon appropriately positioned at the end of the polypeptide to 20 be translated.

In addition, the constructs may contain control regions that regulate as well as engender expression. Generally, in accordance with many commonly practiced procedures, such regions will operate by controlling 25 transcription, such as repressor binding sites and enhancers, among others.

Vectors for propagation and expression generally will include selectable markers. Such markers also may be suitable for amplification or the vectors may contain 30 additional markers for this purpose. In this regard, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells. Preferred markers include dihydrofolate reductase or neomycin resistance for 35 eukaryotic cell culture, and tetracycline or ampicillin

resistance genes for culturing in *E. coli* and other bacteria.

The vector containing the appropriate DNA sequence as described elsewhere herein, as well as an appropriate 5 promoter, and other appropriate control sequences, may be introduced into an appropriate host using a variety of well known techniques suitable to expression therein of a desired polypeptide. Representative examples of appropriate hosts include bacterial cells, such as *E. coli*, 10 *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila S2* and *Spodoptera Sf9* cells; animal cells such as CHO, COS and Bowes melanoma cells; and plant cells. Hosts for a great variety of expression constructs are well known, and 15 those of skill will be enabled by the present disclosure readily to select a host for expressing a LSG polypeptide in accordance with this aspect of the present invention.

More particularly, the present invention also includes recombinant constructs, such as expression 20 constructs, comprising one or more of the sequences described above. The constructs comprise a vector, such as a plasmid or viral vector, into which such LSG sequence of the invention has been inserted. The sequence may be inserted in a forward or reverse orientation. In certain 25 preferred embodiments in this regard, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and there are many commercially available 30 vectors suitable for use in the present invention.

The following vectors, which are commercially available, are provided by way of example. Among vectors preferred for use in bacteria are pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, 35 Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A,

available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are PWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL 5 available from Pharmacia. These vectors are listed solely by way of illustration of the many commercially available and well known vectors that are available to those of skill in the art for use in accordance with this aspect of the present invention. It will be appreciated by those of 10 skill in the art upon reading this disclosure that any other plasmid or vector suitable for introduction, maintenance, propagation and/or expression of a LSG polynucleotide or polypeptide of the invention in a host may be used in this aspect of the invention.

15 Promoter regions can be selected from any desired gene using vectors that contain a reporter transcription unit lacking a promoter region, such as a chloramphenicol acetyl transferase ("cat") transcription unit, downstream of a restriction site or sites for introducing a candidate 20 promoter fragment; i.e., a fragment that may contain a promoter. As is well known, introduction into the vector of a promoter-containing fragment at the restriction site upstream of the cat gene engenders production of CAT activity detectable by standard CAT assays. Vectors 25 suitable to this end are well known and readily available. Two such vectors are pKK232-8 and pCM7. Thus, promoters for expression of LSG polynucleotides of the present invention include, not only well known and readily available promoters, but also promoters that readily may be 30 obtained by the foregoing technique, using a reporter gene.

Among known bacterial promoters suitable for expression of polynucleotides and polypeptides in accordance with the present invention are the *E. coli* lacI and lacZ promoters, the T3 and T7 promoters, the gpt 35 promoter, the lambda PR, PL promoters and the trp promoter.

Among known eukaryotic promoters suitable in this regard are the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous 5 sarcoma virus ("RSV"), and metallothionein promoters, such as the mouse metallothionein-I promoter.

Selection of appropriate vectors and promoters for expression in a host cell is a well known procedure and the requisite techniques for expression vector construction, 10 introduction of the vector into the host and expression in the host are routine skills in the art.

The present invention also relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, 15 or a lower eukaryotic cell, such as a yeast cell. Alternatively, the host cell can be a prokaryotic cell, such as a bacterial cell.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran 20 mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al. *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986).

25 Constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, LSG polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

30 Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

35 Appropriate cloning and expression vectors for use with

prokaryotic and eukaryotic hosts are described by Sambrook et al. cited elsewhere herein.

Generally, recombinant expression vectors will include origins of replication, a promoter derived from a 5 highly-expressed gene to direct transcription of a downstream structural sequence, and a selectable marker to permit isolation of vector containing cells after exposure to the vector. Among suitable promoters are those derived from the genes that encode glycolytic enzymes such as 3-10 phosphoglycerate kinase ("PGK"),  $\alpha$ -factor, acid phosphatase, and heat shock proteins, among others.

Selectable markers include the ampicillin resistance gene of *E. coli* and the *trpL* gene of *S. cerevisiae*.

Transcription of DNA encoding the LSG polypeptides of 15 the present invention by higher eukaryotes may be increased by inserting an enhancer sequence into the vector.

Enhancers are cis-acting elements of DNA, usually about from 10 to 300 base pairs (bp) that act to increase transcriptional activity of a promoter in a given host 20 cell-type. Examples of enhancers include the SV40 enhancer, which is located on the late side of the replication origin at bp 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

25 A polynucleotide of the present invention, encoding a heterologous structural sequence of a LSG polypeptide of the present invention, generally will be inserted into the vector using standard techniques so that it is operably linked to the promoter for expression. The polynucleotide 30 will be positioned so that the transcription start site is located appropriately 5' to a ribosome binding site. The ribosome binding site will be 5' to the AUG that initiates translation of the polypeptide to be expressed. Generally, there will be no other open reading frames that begin with 35 an initiation codon, usually AUG, lying between the

202520 202520 202520 202520 202520

ribosome binding site and the initiating AUG. Also, generally, there will be a translation stop codon at the end of the polypeptide and there will be a polyadenylation signal and a transcription termination signal appropriately disposed at the 3' end of the transcribed region.

Appropriate secretion signals may be incorporated into the expressed polypeptide for secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the 10 extracellular environment. The signals may be endogenous to the polypeptide or they may be heterologous signals.

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous 15 functional regions. Thus, for instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell during purification or during subsequent handling and storage. A 20 region also may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate 25 purification, among others, are familiar and routine techniques in the art.

Suitable prokaryotic hosts for propagation, maintenance or expression of LSG polynucleotides and polypeptides in accordance with the invention include 30 *Escherichia coli*, *Bacillus subtilis* and *Salmonella typhimurium*. Various species of *Pseudomonas*, *Streptomyces*, and *Staphylococcus* are suitable hosts in this regard. Many other hosts also known to those of skill may also be employed in this regard.

DRAFTED 12/20/1990

As a representative, but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising 5 genetic elements of the well known cloning vector pBR322. Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, Wis., USA). These pBR322 "backbone" sections are combined with an appropriate 10 promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, where the selected promoter is inducible it is induced by appropriate means (e.g., temperature shift or exposure to 15 chemical inducer) and cells are cultured for an additional period. Cells typically then are harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification. Microbial cells employed in expression of 20 proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well known to those skilled in the art.

Various mammalian cell culture systems can be 25 employed for expression, as well. An exemplary mammalian expression systems is the COS-7 line of monkey kidney fibroblasts described in Gluzman et al., Cell 23: 175 (1981). Other mammalian cell lines capable of expressing a compatible vector include for example, the C127, 3T3, CHO, 30 HeLa, human kidney 293 and BHK cell lines. Mammalian expression vectors comprise an origin of replication, a suitable promoter and enhancer, and any ribosome binding sites, polyadenylation sites, splice donor and acceptor sites, transcriptional termination sequences, and 5' 35 flanking non-transcribed sequences that are necessary for

expression. In certain preferred embodiments in this regard DNA sequences derived from the SV40 splice sites, and the SV40 polyadenylation sites are used for required non-transcribed genetic elements of these types.

5 LSG polypeptides can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, 10 affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the 15 polypeptide is denatured during isolation and or purification.

LSG polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques 20 from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the LSG polypeptides of the present invention may be glycosylated or may be non- 25 glycosylated. In addition, LSG polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

30 LSG polynucleotides and polypeptides may be used in accordance with the present invention for a variety of applications, particularly those that make use of the chemical and biological properties of the LSGs. Additional applications relate to diagnosis and to treatment of disorders of cells, tissues and organisms. These aspects of

the invention are illustrated further by the following discussion.

**Polynucleotide assays**

As discussed in some detail *supra*, this invention is 5 also related to the use of LSG polynucleotides to detect complementary polynucleotides such as, for example, as a diagnostic reagent. Detection of a mutated form of LSG associated with a dysfunction will provide a diagnostic tool that can add to or define a diagnosis of a disease or 10 susceptibility to a disease which results from under-expression, over-expression or altered expression of a LSG, such as, for example, a susceptibility to inherited lung cancer.

Individuals carrying mutations in a human LSG gene 15 may be detected at the DNA level by a variety of techniques. Nucleic acids for diagnosis may be obtained from a patient's cells, such as from blood, urine, saliva, tissue biopsy and autopsy material. The genomic DNA may be used directly for detection or may be amplified 20 enzymatically using PCR prior to analysis (Saiki et al., *Nature*, 324: 163-166 (1986)). RNA or cDNA may also be used in a similar manner. As an example, PCR primers complementary to a LSG polynucleotide of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 25 19, 20 or 74 can be used to identify and analyze LSG expression and mutations. For example, deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified 30 DNA to radiolabeled LSG RNA or alternatively, radiolabeled LSG antisense DNA sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures.

Sequence differences between a reference gene and 35 genes having mutations also may be revealed by direct DNA

sequencing. In addition, cloned DNA segments may be employed as probes to detect specific DNA segments. The sensitivity of such methods can be greatly enhanced by appropriate use of PCR or another amplification method.

5 For example, a sequencing primer is used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabeled nucleotide or by automatic sequencing procedures with  
10 fluorescent-tags.

Genetic testing based on DNA sequence differences may be achieved by detection of alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents. Small sequence deletions and insertions  
15 can be visualized by high resolution gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific  
20 melting or partial melting temperatures (see, e.g., Myers et al., *Science*, 230: 1242 (1985)).

Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (e.g., Cotton  
25 et al., *Proc. Natl. Acad. Sci., USA*, 85: 4397-4401 (1985)).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes, (e.g., restriction fragment  
30 length polymorphisms ("RFLP") and Southern blotting of genomic DNA. In addition to more conventional gel-electrophoresis and DNA sequencing, mutations also can be detected by *in situ* analysis.

***Chromosome assays***

The LSG sequences of the present invention are also valuable for chromosome identification. There is a need for identifying particular sites on the chromosome and few chromosome marking reagents based on actual sequence data 5 (repeat polymorphisms) are presently available for marking chromosomal location. Each LSG sequence of the present invention is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Thus, the LSGs can be used in the mapping of 10 DNAs to chromosomes, an important first step in correlating sequences with genes associated with disease.

In certain preferred embodiments in this regard, the cDNA herein disclosed is used to clone genomic DNA of a LSG of the present invention. This can be accomplished using a 15 variety of well known techniques and libraries, which generally are available commercially. The genomic DNA is used for *in situ* chromosome mapping using well known techniques for this purpose.

In some cases, sequences can be mapped to chromosomes 20 by preparing PCR primers (preferably 15-25 bp) from the cDNA. Computer analysis of the 3' untranslated region of the gene is used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers are then used for 25 PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the primer will yield an amplified fragment.

PCR mapping of somatic cell hybrids is a rapid 30 procedure for assigning a particular DNA to a particular chromosome. Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes or pools of large genomic clones in an analogous manner. Other 35 mapping strategies that can similarly be used to map to its

chromosome include *in situ* hybridization, prescreening with labeled flow-sorted chromosomes and preselection by hybridization to construct chromosome specific-cDNA libraries.

5       Fluorescence *in situ* hybridization ("FISH") of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with cDNA as short as 50 or 60 bp. This technique is described by Verma et al. (HUMAN  
10 CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES, Pergamon Press, New York (1988)).

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data.

15 Such data are found, for example, in V. McKusick, MENDELIAN INHERITANCE IN MAN, available on line through Johns Hopkins University, Welch Medical Library. The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage  
20 analysis (coinheritance of physically adjacent genes).

Next, it is necessary to determine the differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal  
25 individuals, then the mutation is likely to be the causative agent of the disease.

With current resolution of physical mapping and genetic mapping techniques, a cDNA precisely localized to a chromosomal region associated with the disease could be one  
30 of between 50 and 500 potential causative genes. (This assumes 1 megabase mapping resolution and one gene per 20 kb).

#### ***Polypeptide assays***

As described in some detail *supra*, the present  
35 invention also relates to diagnostic assays such as

PENDING PCT/US88/0458

quantitative and diagnostic assays for detecting levels of LSG polypeptide in cells and tissues, and biological fluids such as blood and urine, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay 5 in accordance with the present invention for detecting over-expression or under-expression of a LSG polypeptide compared to normal control tissue samples may be used to detect the presence of neoplasia. Assay techniques that can be used to determine levels of a protein, such as a LSG 10 polypeptide of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays. Among these ELISAs frequently are preferred.

15 For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10  $\mu$ g/ml. The antibodies are either monoclonal or polyclonal 20 and are produced by methods as described herein. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced. The coated wells are then incubated for > 2 hours at room temperature with a sample containing the LSG polypeptide. Preferably, serial 25 dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide. Next, 50  $\mu$ l of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 30 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate. 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution (75 $\mu$ l) is then added to each well and the plate is incubated 1 35 hour at room temperature. The reaction is measured by a

microtiter plate reader. A standard curve is prepared using serial dilutions of a control sample, and polypeptide concentration is plotted on the X-axis (log scale) while fluorescence or absorbance is plotted on the Y-axis (linear scale). The concentration of the LSG polypeptide in the sample is interpolated using the standard curve.

**Antibodies**

As discussed in some detail *supra*, LSG polypeptides, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be polyclonal or monoclonal antibodies. The present invention also includes chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

A variety of methods for antibody production are set forth in Current Protocols, Chapter 2.

For example, cells expressing a LSG polypeptide of the present invention can be administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. This preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity. The antibody obtained will bind with the LSG polypeptide itself. In this manner, even a sequence encoding only a fragment of the LSG polypeptide can be used to generate antibodies binding the whole native polypeptide. Such antibodies can then be used to isolate the LSG polypeptide from tissue expressing that LSG polypeptide.

Alternatively, monoclonal antibodies can be prepared. Examples of techniques for production of monoclonal

0230-2420-2000

antibodies include, but are not limited to, the hybridoma technique (Kohler, G. and Milstein, C., *Nature* 256: 495-497 (1975), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., *Immunology Today* 4: 72 (1983) and 5 (Cole et al., pg. 77-96 in *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc. (1985). The EBV-hybridoma technique is useful in production of human monoclonal antibodies.

Hybridoma technologies have also been described by 10 Khler et al. (*Eur. J. Immunol.* 6: 511 (1976)) Khler et al. (*Eur. J. Immunol.* 6: 292 (1976)) and Hammerling et al. (in: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N. Y., pp. 563-681 (1981)). In general, such procedures involve immunizing an animal (preferably a mouse) with LSG 15 polypeptide or, more preferably, with a secreted LSG polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum 20 (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin. The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line 25 may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP20), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described 30 by Wands et al. (*Gastroenterology* 80: 225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of 35 binding to the polypeptide can be produced in a two-step

procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance 5 with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the 10 protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

15 Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can also be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention. Also, transgenic mice, as well as other nonhuman transgenic animals, may be 20 used to express humanized antibodies to immunogenic polypeptide products of this invention.

It will be appreciated that Fab, F(ab')<sub>2</sub> and other fragments of the antibodies of the present invention may also be used according to the methods disclosed herein.

25 Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')<sub>2</sub> fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA 30 technology or through synthetic chemistry.

For *in vivo* use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the 35 monoclonal antibodies described above. Methods for

producing chimeric antibodies are known in the art (See, for review, Morrison, Science 229: 1202 (1985); Oi et al., BioTechniques 4: 214 (1986); Cabilly et al., U. S. Patent 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 5 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312: 643 (1984); Neuberger et al., Nature 314: 268 (1985).)

The above-described antibodies may be employed to isolate or to identify clones expressing LSG polypeptides 10 or purify LSG polypeptides of the present invention by attachment of the antibody to a solid support for isolation and/or purification by affinity chromatography. As discussed in more detail *supra*, antibodies specific against a LSG may also be used to image tumors, particularly cancer 15 of the lung, in patients suffering from cancer. Such antibodies may also be used therapeutically to target tumors expressing a LSG.

Preferred exemplary antigenic epitopes of LSGs of the present invention which have been identified are depicted 20 below. The antigenicity index (AI avg) used is Jameson-Wolf. In some embodiment, it may be preferred to raise antibodies against these regions of the LSGs.

DEX73\_2.aa        Antigenicity Index(Jameson-Wolf)  
(SEQ ID NO:75)

25        positions        AI avg        length  
          16-50            1.06            35

DEX73\_3.aa        Antigenicity Index(Jameson-Wolf)  
(SEQ ID NO:76)

30        positions        AI avg        length  
          52-66            1.05            15

DEX73\_5.aa        Antigenicity Index(Jameson-Wolf)  
(SEQ ID NO:77)

35        positions        AI avg        length  
          1419-1433        1.16            15  
          1387-1414        1.08            28

808-825 1.00 18  
DEX73\_8.aa Antigenicity Index (Jameson-Wolf)  
(SEQ ID NO:78)

	positions	AI avg	length
5	208-223	1.06	16
	123-135	1.05	13
	689-717	1.03	29
	63-90	1.02	28
	653-683	1.01	31
10	366-377	1.00	12

DEX73\_12.aa Antigenicity Index (Jameson-Wolf)  
(SEQ ID NO:80)

	positions	AI avg	length
15	56-68	1.05	13

DEX73\_13.aa Antigenicity Index (Jameson-Wolf)  
(SEQ ID NO:81)

	positions	AI avg	length
20	207-217	1.28	11
	72-85	1.17	14
20	405-469	1.15	65
	151-171	1.02	21

DEX73\_18.aa Antigenicity Index (Jameson-Wolf)  
(SEQ ID NO:84)

	positions	AI avg	length
25	40-51	1.27	12

#### *LSG binding molecules and assays*

This invention also provides a method for identification of molecules, such as receptor molecules, 30 that bind LSGs. Genes encoding proteins that bind LSGs, such as receptor proteins, can be identified by numerous methods known to those of skill in the art. Examples include, but are not limited to, ligand panning and FACS sorting. Such methods are described in many laboratory

manuals such as, for instance, Coligan et al., Current Protocols in Immunology 1(2): Chapter 5 (1991).

Expression cloning may also be employed for this purpose. To this end, polyadenylated RNA is prepared from 5 a cell responsive to a LSG of the present invention. A cDNA library is created from this RNA and the library is divided into pools. The pools are then transfected individually into cells that are not responsive to a LSG of the present invention. The transfected cells then are exposed to 10 labeled LSG. LSG polypeptides can be labeled by a variety of well-known techniques including, but not limited to, standard methods of radio-iodination or inclusion of a recognition site for a site-specific protein kinase. Following exposure, the cells are fixed and binding of 15 labeled LSG is determined. These procedures conveniently are carried out on glass slides. Pools containing labeled LSG are identified as containing cDNA that produced LSG-binding cells. Sub-pools are then prepared from these positives, transfected into host cells and screened as 20 described above. Using an iterative sub-pooling and re-screening process, one or more single clones that encode the putative binding molecule, such as a receptor molecule, can be isolated.

Alternatively a labeled ligand can be photoaffinity 25 linked to a cell extract, such as a membrane or a membrane extract, prepared from cells that express a molecule that it binds, such as a receptor molecule. Cross-linked material is resolved by polyacrylamide gel electrophoresis ("PAGE") and exposed to X-ray film. The labeled complex 30 containing the ligand-receptor can be excised, resolved into peptide fragments, and subjected to protein microsequencing. The amino acid sequence obtained from microsequencing can be used to design unique or degenerate oligonucleotide probes to screen cDNA libraries to identify 35 genes encoding the putative receptor molecule.

20250000-20250000

Polypeptides of the invention also can be used to assess LSG binding capacity of LSG binding molecules, such as receptor molecules, in cells or in cell-free preparations.

5 **Agonists and antagonists - assays and molecules**

The invention also provides a method of screening compounds to identify those which enhance or block the action of a LSG on cells. By "compound", as used herein, it is meant to be inclusive of small organic molecules, 10 peptides, polypeptides and antibodies as well as any other candidate molecules which have the potential to enhance or agonize or block or antagonize the action of LSG on cells.

As used herein, an agonist is a compound which increases the natural biological functions of a LSG or which

15 functions in a manner similar to a LSG, while an antagonist, as used herein, is a compound which decreases or eliminates such functions. Various known methods for screening for agonists and/or antagonists can be adapted for use in identifying LSG agonist or antagonists.

20 For example, a cellular compartment, such as a membrane or a preparation thereof, such as a membrane-preparation, may be prepared from a cell that expresses a molecule that binds a LSG, such as a molecule of a signaling or regulatory pathway modulated by LSG. The 25 preparation is incubated with labeled LSG in the absence or the presence of a compound which may be a LSG agonist or antagonist. The ability of the compound to bind the binding molecule is reflected in decreased binding of the labeled ligand. Compounds which bind gratuitously, i.e., 30 without inducing the effects of a LSG upon binding to the LSG binding molecule are most likely to be good antagonists. Compounds that bind well and elicit effects that are the same as or closely related to LSG are agonists. LSG-like effects of potential agonists and 35 antagonists may be measured, for instance, by determining

activity of a second messenger system following interaction of the candidate molecule with a cell or appropriate cell preparation, and comparing the effect with that of LSG or molecules that elicit the same effects as LSG. Second messenger systems that may be useful in this regard include, but are not limited to, AMP guanylate cyclase, ion channel or phosphoinositide hydrolysis second messenger systems.

Another example of an assay for LSG antagonists is a competitive assay that combines LSG and a potential antagonist with membrane-bound LSG receptor molecules or recombinant LSG receptor molecules under appropriate conditions for a competitive inhibition assay. LSG can be labeled, such as by radioactivity, such that the number of LSG molecules bound to a receptor molecule can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a LSG polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a receptor molecule, without inducing LSG-induced activities, thereby preventing the action of LSG by excluding LSG from binding.

Potential antagonists include small molecules which bind to and occupy the binding site of the LSG polypeptide thereby preventing binding to cellular binding molecules, such as receptor molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules.

Other potential antagonists include antisense molecules. Antisense technology can be used to control gene

expression through antisense DNA or RNA or through triple-helix formation. Antisense techniques are discussed, for example, in Okano, J. Neurochem. 56: 560 (1991); OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION, CRC Press, Boca Raton, Fla. (1988). Triple helix formation is discussed in, for instance Lee et al., Nucleic Acids Research 6: 3073 (1979); Cooney et al., Science 241: 456 (1988); and Dervan et al., Science 251: 1360 (1991). The methods are based on binding of a polynucleotide to a complementary DNA or RNA. For example, the 5' coding portion of a polynucleotide that encodes a mature LSG polypeptide of the present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of a LSG polypeptide. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into a LSG polypeptide. The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of a LSG.

#### ***Compositions***

The present invention also relates to compositions comprising a LSG polynucleotide or a LSG polypeptide or an agonist or antagonist thereof.

For example, a LSG polynucleotide, polypeptide or an agonist or antagonist thereof of the present invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide of the invention and a pharmaceutically acceptable carrier or excipient. Such

SEARCHED  
INDEXED  
MAILED

carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration.

5 Compositions of the present invention will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the polypeptide or other compound 10 alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically 15 effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1,  $\mu\text{g}/\text{kg}/\text{day}$  to 10  $\text{mg}/\text{kg}/\text{day}$  of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 20  $\text{mg}/\text{kg}/\text{day}$ , and most preferably for humans between about 0.01 and 1  $\text{mg}/\text{kg}/\text{day}$  for the hormone. If given continuously, the polypeptide or other compound is typically administered at a dose rate of about 1  $\mu\text{g}/\text{kg}/\text{hour}$  to about 50  $\text{mg}/\text{kg}/\text{hour}$ , either by 1-4 injections per day or 25 by continuous subcutaneous infusion, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

30 Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracistemally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), bucally, or as an oral 35 or nasal spray. "Pharmaceutically acceptable carrier"

DEPARTMENT OF COMMERCE

refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, 5 intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The polypeptide or other compound is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include 10 semipermeable polymer matrices in the form of shaped articles, e. g., films, or microcapsules. Sustained-release matrices include polylactides (U.S. Patent 3,773,919 and EP 58481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., *Biopolymers* 22: 547-556 15 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., *J. Biomed. Mater. Res.* 15: 167-277 (1981), and R. Langer, *Chem. Tech.* 12: 98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) and poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also 20 include liposomally entrapped polypeptides. Liposomes containing the polypeptide or other compound are prepared by well known methods (Epstein et al., *Proc. Natl. Acad. Sci. USA* 82: 3688-3692 (1985); Hwang et al., *Proc. Natl. Acad. Sci. USA* 77: 4030-4034 (1980); EP 52322; EP 36676; EP 25 88046; EP 143949; EP 142641; Japanese Pat. Appl. 83-118008; U.S. Patent 4,485,045 and 4,544,545; and EP 102324). Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the 30 selected proportion being adjusted for the optimal therapy.

For parenteral administration, in one embodiment, the polypeptide or other compound is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a 35 pharmaceutically acceptable carrier, i.e., one that is non-

toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation.

For example, the formulation preferably does not 5 include oxidizing agents and other compounds that are known to be deleterious to the polypeptide or other compound.

Generally, the formulations are prepared by 10 contacting the polypeptide or other compound uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped 15 into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, 20 and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of 25 additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e. g., 30 polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose 35 or its derivatives, glucose, mannose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or 35 PEG.

P0230-00000000000000000000000000000000

The polypeptide or other compound is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of 5 the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts or salts of the other compounds.

Any polypeptide to be used for therapeutic administration should be sterile. Sterility is readily 10 accomplished by filtration through sterile filtration membranes (e. g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper 15 pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a 20 lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1 % (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

25 **Kits**

The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention. Associated with such 30 container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, reflecting approval by the agency of the manufacture, use or sale of the product for human administration.

35 **Administration**

LSG polypeptides or polynucleotides or other compounds, preferably agonists or antagonists thereof of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic 5 compounds.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, 10 intranasal or intradermal routes among others.

The pharmaceutical compositions generally are administered in an amount effective for treatment or prophylaxis of a specific indication or indications. In general, the compositions are administered in an amount of 15 at least about 10  $\mu\text{g}/\text{kg}$  body weight. However, it will be appreciated that optimum dosage will be determined by standard methods for each treatment modality and indication, taking into account the indication, its severity, route of administration, complicating conditions 20 and the like.

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a LSG polypeptide in an individual can be treated by administering the LSG polypeptide of the present invention, 25 preferably in the secreted form, or an agonist thereof. Thus, the invention also provides a method of treatment of an individual in need of an increased level of a LSG polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the 30 LSG polypeptide or an agonist thereof to increase the activity level of the LSG polypeptide in such an individual. For example, a patient with decreased levels of a LSG polypeptide may receive a daily dose 0.1-100  $\mu\text{g}/\text{kg}$  of a LSG polypeptide or agonist thereof for six consecutive

SEARCHED - INDEXED - RECORDED - SERIALIZED - FILED

days. Preferably, if a LSG polypeptide is administered it is in the secreted form.

Compositions of the present invention can also be administered to treating increased levels of a LSG polypeptide. For example, antisense technology can be used to inhibit production of a LSG polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer. A patient diagnosed with abnormally increased levels of a polypeptide can be administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is preferably repeated after a 7-day rest period if the treatment was well tolerated.

Compositions comprising an antagonist of a LSG polypeptide can also be administered to decrease levels of LSG in a patient.

#### ***Gene therapy***

The LSG polynucleotides, polypeptides, agonists and antagonists that are polypeptides may be employed in accordance with the present invention by expression of such polypeptides *in vivo*, in treatment modalities often referred to as "gene therapy."

Thus, for example, cells from a patient may be engineered with a polynucleotide, such as a DNA or RNA, encoding a polypeptide *ex vivo*, and the engineered cells then can be provided to a patient to be treated with the polypeptide. For example, cells may be engineered *ex vivo* by the use of a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention. Such methods are well-known in the art and their use in the present invention will be apparent from the teachings herein.

Similarly, cells may be engineered *in vivo* for expression of a polypeptide *in vivo* by procedures known in

the art. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed *supra*. The retroviral expression construct then may be isolated and introduced 5 into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a patient for 10 engineering cells *in vivo* and expression of the polypeptide *in vivo*. These and other methods for administering a polypeptide of the present invention would be apparent to those skilled in the art upon reading the instant application.

15            Retroviruses from which the retroviral plasmid  
vectors herein above mentioned may be derived include, but  
are not limited to, Moloney Murine Leukemia Virus, spleen  
necrosis virus, retroviruses such as Rous Sarcoma Virus,  
Harvey Sarcoma Virus, avian leukosis virus, gibbon ape  
20          leukemia virus, human immunodeficiency virus, adenovirus,  
Myeloproliferative Sarcoma Virus, and mammary tumor virus.  
In one embodiment, the retroviral plasmid vector is derived  
from Moloney Murine Leukemia Virus.

Such vectors will include one or more promoters for expressing the polypeptide. The selection of a suitable promoter will be apparent to those skilled in the art from the teachings contained herein. However, examples of suitable promoters which may be employed include, but are not limited to, the retroviral LTR, the SV40 promoter, the human cytomegalovirus (CMV) promoter described in Miller et al., Biotechniques 7: 980-990 (1989), and eukaryotic cellular promoters such as the histone, RNA polymerase III, and beta-actin promoters. Other viral promoters which may be employed include, but are not limited to, adenovirus promoters, thymidine kinase (TK) promoters, and B19

parvovirus promoters. Additional promoters which may be used include respiratory syncytial virus (RSV) promoter, inducible promoters such as the MMT promoter, the metallothionein promoter, heat shock promoters, the albumin 5 promoter, the ApoAI promoter, human globin promoters, viral thymidine kinase promoters such as the Herpes Simplex thymidine kinase promoter, retroviral LTRs, the beta-actin promoter, and human growth hormone promoters. The promoter also may be the native promoter which controls the gene 10 encoding the polypeptide.

The nucleic acid sequence encoding the polypeptide of the present invention will be placed under the control of a suitable promoter.

In one embodiment, the retroviral plasmid vector is 15 employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the PE501, PA317, Y-2, Y-AM, PA12, T19-14X, VT-19-17-H2, YCRE, YCRIP, GP+E-86, GP+envAml2, and DAN cell lines as described in 20 Miller, A., Human Gene Therapy 1: 5-14 (1990). The vector may be transduced into the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation, the use of liposomes, and  $\text{CaPO}_4$  precipitation. Alternatively, the retroviral plasmid 25 vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host. The producer cell line will generate infectious retroviral vector particles which are inclusive of the nucleic acid sequence(s) 30 encoding the polypeptides. Such retroviral vector particles then may be employed to transduce eukaryotic cells, either *in vitro* or *in vivo*. The transduced eukaryotic cells will express the nucleic acid sequence(s) 35 encoding the polypeptide. Eukaryotic cells which may be transduced include, but are not limited to, embryonic stem cells, embryonic carcinoma cells, as well as hematopoietic

stem cells, hepatocytes, fibroblasts, myoblasts, keratinocytes, endothelial cells, and bronchial epithelial cells.

An exemplary method of gene therapy involves 5 transplantation of fibroblasts which are capable of expressing a LSG polypeptide or an agonist or antagonist thereof onto a patient. Generally fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small 10 pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night.

After 24 hours at room temperature, the flask is inverted 15 and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e. g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week. At this time, fresh media is added and subsequently changed every 20 several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks. pMV-7 (Kirschmeier, P. T. et al., DNA, 7: 219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma 25 virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads. The cDNA encoding a LSG polypeptide of the present invention or an agonist or antagonist thereof can 30 be amplified using PCR primers which correspond to their 5' and 3' end sequences respectively. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and 35 HindIII fragment are added together in the presence of T4

DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB 101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted. Amphotropic pA317 or GP+aml2 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells). Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced. The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

Alternatively, *in vivo* gene therapy methods can be used to treat LSG related disorders, diseases and conditions. Gene therapy methods relate to the

introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide.

For example, a LSG polynucleotide of the present

5 invention or a nucleic acid sequence encoding an agonist or antagonist thereto may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art,  
10 see, for example, WO 90/11092, WO 98/11779; U.S. Patents 5,693,622, 5,705,151, and 5,580,859; Tabata H. et al. (1997) *Cardiovasc. Res.* 35 (3): 470-479, Chao J et al. (1997) *Pharmacol. Res.* 35 (6): 517-522, Wolff J. A. (1997) *Neuromuscul. Disord.* 7 (5): 314-318, Schwartz B. et al.  
15 (1996) *Gene Ther.* 3 (5): 405-411, Tsurumi Y. et al. (1996) *Circulation* 94 (12): 3281-3290 (incorporated herein by reference). The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the  
20 interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers  
25 to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, polynucleotides may also be delivered in  
30 liposome formulations (such as those taught in Felgner P. L. et al. (1995) *Ann. NY Acad. Sci.* 772: 126-139 and Abdallah B. et al. (1995) *Biol. Cell* 85 (1): 1-7) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells.

Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal,

15 including of muscle, skin, brain, lung, liver, spleen, bone  
marrow, thymus, heart, lymph, blood, bone, cartilage,  
pancreas, kidney, gall bladder, stomach, intestine, testis,  
ovary, uterus, rectum, nervous system, eye, gland, and  
connective tissue. Interstitial space of the tissues  
20 comprises the intercellular fluid, mucopolysaccharide  
matrix among the reticular fibers of organ tissues, elastic  
fibers in the walls of vessels or chambers, collagen fibers  
of fibrous tissues, or that same matrix within connective  
tissue ensheathing muscle cells or in the lacunae of bone.  
25 It is similarly the space occupied by the plasma of the  
circulation and the lymph fluid of the lymphatic channels.  
Delivery to the interstitial space of muscle tissue is  
preferred. The polynucleotide construct may be  
conveniently delivered by injection into the tissues  
30 comprising these cells. They are preferably delivered to  
and expressed in persistent, non-dividing cells which are  
differentiated, although delivery and expression may be  
achieved in non-differentiated or less completely  
differentiated cells, such as, for example, stem cells of  
35 blood or skin fibroblasts. *In vivo* muscle cells are

particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from 5 about 0.05  $\mu$ g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the 10 tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral 15 route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide 20 constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle *in vivo* is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide 25 of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of 30 the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. 35 The template DNA is injected in 0.1 ml of carrier in a 1 cc

syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and 5 the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15  $\mu$ m cross-section of the individual quadriceps muscles is histochemically stained 10 for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing 15 total cellular DNA and HIRT supernatants from injected and control mice.

The results of the above experimentation in mice can be used to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

#### 20 *Nonhuman Transgenic Animals*

The LSG polypeptides of the invention can also be expressed in nonhuman transgenic animals. Nonhuman animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, 25 sheep, cows and non-human primates, e. g., baboons, monkeys, and chimpanzees, may be used to generate transgenic animals. Any technique known in the art may be used to introduce the transgene (I. e., polynucleotides of the invention) into animals to produce the founder lines of 30 transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., Appl. Microbiol. Biotechnol. 40: 691-698 (1994); Carver et al., Biotechnology (NY) 11: 1263-1270 (1993); Wright et al., Biotechnology (NY) 9: 830-834 (1991); and Hoppe et 35 al., U.S. Patent 4,873,191); retrovirus mediated gene

transfer into germ lines (Van der Putten et al., Proc. Natl. Acad. Sci., USA 82: 6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., Cell 56: 313-321 (1989)); electroporation of cells 5 or embryos (Lo, 1983, Mol. Cell. Biol. 3: 1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e.g., Ulmer et al., Science 259: 1745 (1993)); introducing nucleic acid constructs into embryonic pluripotent stem cells and 10 transferring the stem cells back into the blastocyst; and sperm mediated gene transfer (Lavitrano et al., Cell 57: 717-723 (1989)). For a review of such techniques, see Gordon, "Transgenic Animals," Intl. Rev. Cytol. 115: 171-229 (1989), which is incorporated by reference herein in its 15 entirety.

Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult 20 cells induced to quiescence (Campell et al., Nature 380: 64-66 (1996); Wilmut et al., Nature 385: 810813 (1997)).

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all 25 their cells, i.e., mosaic or chimeric animals. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, e. g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular 30 cell type by following, for example, the teaching of Lasko et al. (Lasko et al., Proc. Natl. Acad. Sci. USA 89: 6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to 35 those of skill in the art. When it is desired that the

polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the 5 endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell 10 type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Science 265: 103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of 15 interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished 20 by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not 25 limited to, Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and reverse transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using 30 antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of 35 founder animals with more than one integration site in

order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of 5 heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; 10 and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems 15 useful in elaborating the biological function of LSG polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression of LSGs, and in screening for compounds effective in ameliorating such LSG associated conditions and/or 20 disorders.

***Knock-Out Animals***

Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination (e. g., see 25 Smithies et al., *Nature* 317: 230-234 (1985); Thomas & Capecchi, *Cell* 51: 503512 (1987); Thompson et al., *Cell* 5: 313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional LSG polynucleotide of the invention (or a 30 completely unrelated DNA sequence) flanked by DNA homologous to the endogenous LSG polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells 35 that express polypeptides of the invention *in vivo*. In

another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results 5 in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (e. g., see Thomas & Capecchi 1987 and Thompson 1989, 10 *supra*). This approach can also be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors that will be apparent to those of skill in the art.

15 In further embodiments of the invention, cells that are genetically engineered to express the LSG polypeptides of the invention, or alternatively, that are genetically engineered not to express the LSG polypeptides of the invention (e. g., knockouts) are administered to a patient 20 *in vivo*. Such cells may be obtained from the patient or a MHC compatible donor and can include, but are not limited to, fibroblasts, bone marrow cells, blood cells (e. g., lymphocytes), adipocytes, muscle cells, and endothelial cells. The cells are genetically engineered *in vitro* using 25 recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e. g., by transduction 30 (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc.

P0230-A200000000000

The coding sequence of the LSG polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the LSG polypeptides of the invention. The engineered cells which express and preferably secrete the LSG polypeptides of the invention can be introduced into the patient systemically, e.g., in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e.g., genetically engineered fibroblasts can be implanted as part of a skin graft or genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft (see, for example, U.S. Patent 5,399,349 and U.S. Patent 5,460,959 each of which is incorporated by reference herein in its entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of LSG polypeptides of the present invention, studying conditions and/or disorders associated with aberrant LSG expression, and in screening for compounds effective in ameliorating such LSG associated conditions and/or disorders.

The following nonlimiting example is provided to further illustrate the present invention.

**EXAMPLE**

The following Example is carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in 5 detail. Routine molecular biology techniques of the following example can be carried out as described in standard laboratory manuals, such as Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. 10 (1989).

**Introduction and background for Microarray analysis**

CDNA microarrays are prepared by high-speed robotic printing of thousands of distinct cDNAs in an ordered array on glass microscope slides. They are used to measure the 15 relative abundance of specific sequences in two complex samples (Schena et al, 1995; Shalon et al, 1996).

In the microarray procedure, mRNA is isolated from tissues of interest, either from a tumor or control (normal or normal adjacent tissue). mRNA (200-600 ng) from cancer 20 tissue or control is reverse transcribed to incorporate the fluorescent nucleotides Cy5 (red) or Cy3 (green), respectively. The two populations of fluorescently labeled cDNA are mixed together and hybridized simultaneously to a microarray bearing approximately 10,000 cDNA elements in a 25 2cm x 2cm area on a glass slide (Microarrays hybridization service: Incyte Genomics, Fremont, CA, USA). After hybridization, the slides are scanned with a scanning laser confocal microscope.

The scanned image is used to generate the intensity 30 and local background measurements for each spot on the array (GEMtools software, Incyte Genomics). For each spot, representing one EST, the ratio of the normalized Cy5/Cy3 intensities generates a quantitation of the gene's expression in one tissue relative to the control, in this

case, the expression in cancer tissue versus either normal or normal adjacent tissue. For example, a gene that shows a Cancer-Cy5 intensity of 3000 and a Normal-Cy3 intensity of 1000 is expressed 3-fold more in cancer tissue.

5 Advanced analysis software is used to sort and decipher patterns of gene expression from the data (Cluster and Treeview programs, Stanford University; Eisen et al, 1998; Alizadeh et al, 2000). However, the reproducibility study from Incyte shows that the level of detectable differential 10 expression is calculated to be approximately plus or minus 1.74. Consequently, any elements with observed ratios greater than or equal to 1.8 between cancer and normal are deemed differentially expressed.

References:

15 1. Schena, M., D. Shalon, R.W. Davis, and P.O. Brown. 1995. Quantitative monitoring of gene expression patterns with a complementary cDNA microarray. *Science* 270: 467-470.

2. Shalon, D., S.J. Smith, and P.O. Brown. 1996. A DNA 20 Microarray System for Analyzing Complex DNA samples Using Two-color Fluorescent Probe Hybridization. *Genome Research* 6: 639-645.

3. Eisen, M.B., P.T. Spellman, P.O. Brown, and D. Botstein. 1998. "Cluster analysis and display of genome-wide 25 expression patterns". *PNAS* 95: 14863-14868.

4. Alizadeh, A.A., et al, 2000. "Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling." *Nature*, 403: 503-511.

5. GEM Microarray Reproducibility Study. Technical 30 specifications from Incyte Genomics.

**Lung diaDexus microarray candidates**

Following is a list of "diaDexus microarray candidates" sequences for lung cancer, also referred to herein as lung specific genes or LSGs:

SEQ ID	Gene ID	ddx lung code	ddx QPCR lung code
1/19	979057	Lng128	Lng128
2/20	347842	Lng129	Lng129
3	983439	Lng112	Lng112
5 4	236582	Lng114	Lng114
5	210995	Lng118	Lng118
6	208994	Lng121	Lng121
7	1066498	Lng124	Lng124
8	287016	Lng126	Lng126
10 9	10717	SQLng001	Lng136
10	24945	SQLng006	Lng143
11	52017	SQLng007	Lng144
12	460254	SQLng110	Lng138
13/74	179090	SQLng012	Lng137
15 14	6348	SQLng004	Lng142
15	94694	SQLng005	Lng140
16	145812	SQLng008	Lng151
17	10713	SQLng002	Lng150
18	20152	SQLng003	Lng141

20 **Example 1**

## Sequence 1

**Lng128**

Gene ID 979057

Table 1. The absolute numbers are relative levels of expression of Lng128 in 24 normal different tissues. All the values are compared to normal trachea (calibrator). These RNA samples are commercially pools, originated by pooling samples of a particular tissue from different individuals.

Tissue	NORMAL
Adrenal Gland	0.03
Bladder	0.00
Brain	6.68
Cervix	0.00
Colon	0.00
Endometrium	0.12
Esophagus	0.00
Heart	0.01

	Kidney	0.02
	Liver	0.03
	Lung	35.63
	Mammary Gland	0.02
5	Muscle	0.00
	Ovary	1.11
	Pancreas	17.94
	Prostate	0.42
	Rectum	0.16
10	Small Intestine	0.00
	Spleen	1.27
	Stomach	0.00
	Testis	2.17
15	Thymus	0.13
	Trachea	1.00
	Uterus	0.09

0=negative

The relative levels of expression in Table 1 show that Lng128 mRNA expression is much higher in lung (35.63) compared with most other normal tissues analyzed.

The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

**Table 2.** The absolute numbers are relative levels of expression of Lng128 in 69 pairs of matching samples and 1 ovary normal and one ovary cancer sample. All the values are compared to normal trachea (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING	
					NORMAL	
					ADJACENT	
35	Lng 60L	Adenocarcinoma	Lung 1	196.04	17.03	
	Lng 143L	Adenocarcinoma	Lung 2	21.93	0.88	
	Lng 60XL	Adenocarcinoma	Lung 3	17.21	42.37	
	Lng AC82	Adenocarcinoma	Lung 4	43.26	4.56	
	Lng AC88	Adenocarcinoma	Lung 5	364.56	101.48	
	Lng AC66	Adenocarcinoma	Lung 6	17.94	13.27	
40	Lng AC69	Adenocarcinoma	Lung 7	582.05	39.12	
	Lng AC11	Adenocarcinoma	Lung 8	24.42	113.38	
	Lng AC32	Adenocarcinoma	Lung 9	648.07	27.19	
	Lng AC39	Adenocarcinoma	Lung 10	249.00	1.71	
	Lng AC94	Adenocarcinoma	Lung 11	42.81	112.99	
45	Lng AC90	Adenocarcinoma	Lung 12	196.72	0.48	
	Lng 47XQ	Adenocarcinoma	Lung 13	88.95	0.54	

	Lng 223L	Adenocarcinoma	Lung 14	5.80	0.00
	Lng 528L	Adenocarcinoma	Lung 15	45.25	177.91
	Lng BR26	Bronchio-alveolar carcinoma	Lung 16	2.80	28.54
	Lng BA641	Bronchogenic carcinoma	Lung 17	1746.20	36.13
5	Lng 315L	Squamous cell carcinoma	Lung 18	1.67	736.73
	Lng SQ45	Squamous cell carcinoma	Lung 19	828.87	62.68
	Lng SQ14	Squamous cell carcinoma	Lung 20	0.07	15.56
	Lng SQ9X	Squamous cell carcinoma	Lung 21	73.26	4.32
	Lng SQ56	Squamous cell carcinoma	Lung 22	33.24	141.53
10	Lng SQ80	Squamous cell carcinoma	Lung 23	101.13	44.79
	Lng SQ32	Squamous cell carcinoma	Lung 24	119.43	9.82
	Lng SQ16	Squamous cell carcinoma	Lung 25	64.00	10.85
	Lng SQ79	Squamous cell carcinoma	Lung 26	52.16	142.52
	Lng 90X	Squamous cell carcinoma	Lung 27	38.72	6.23
15	Lng BR94	Squamous cell carcinoma	Lung 28	27.19	0.00
	Lng C20X	Squamous cell carcinoma	Lung 29	0.00	1.59
	Lng SQ44	Squamous cell carcinoma	Lung 30	13.88	0.04
	Lng SQ43	Squamous cell carcinoma	Lung 31	24.00	1.39
20	Lng 77L	Large cell carcinoma	Lung 32	0.15	13.93
	Lng LC71	Large cell carcinoma	Lung 33	61.61	190.68
	Lng LC109	Large cell carcinoma	Lung 34	25.19	513.78
	Lng LC80	Large cell carcinoma	Lung 35	537.45	47.01
	Lng 75XC	Metastatic from bone cancer	Lung 36	44.79	39.95
	Lng MT71	Metastatic from renal cell cancer	Lung 37	11.35	26.45
25	Lng MT67	Metastatic from melanoma	Lung 38	3.28	7.97
	Bld 46XK		Bladder 1	0.00	0.00
	Bld TR14		Bladder 2	0.46	0.00
	Cvx KS52		Cervix 1	0.29	0.00
	Cvx KS83		Cervix 2	0.00	0.00
30	Cln AS45		Colon 1	0.00	0.00
	Cln RC01		Colon 2	0.00	0.10
	End 8911		Endometrium 1	0.08	0.68
	End 28XA		Endometrium 2	12.73	0.57
35	kid 107XD		Kidney 1	0.02	0.02
	Kid 109XD		Kidney 2	0.05	0.25
	Liv 94XA		Liver 1	0.00	0.00
	Liv 174L		Liver 2	0.00	0.00
	Mam 162X		Mammary 1	0.00	0.02
40	Mam 497M		Mammary 2	0.00	0.00
	Ovr A082		Ovary 1	0.03	1.57
	Ovr 18GA		Ovary 2		5.78
	Ovr 180B		Ovary 3	0.03	

	Pan 71X	Pancreas 1	0.03	0.02
	Pan 92X	Pancreas 2	0.65	0.00
	Pro	Prostate 1	0.01	0.03
	109XB			
5	Pro	Prostate 2	0.02	0.02
	125XB			
	Skn 248S	Skin 1	0.11	0.00
	Skn 816S	Skin 2	1.01	0.00
	SmInt	Small	0.01	0.00
10	21XA	Intestine 1		
	SmInt H89	Small	0.67	2.76
		Intestine 2		
	Sto 758S	Stomach 1	0.00	0.00
	Sto 531S	Stomach 2	0.08	0.00
	Tst 647T	Testis 1	4.38	0.96
15	Tst 39X	Testis 2	8.69	1.19
	Thr 143N	Thyroid 1	0.15	0.00
	Thr 270T	Thyroid 2	0.00	0.00
	Utr 135XO	Uterus 1	0.19	0.27
	Utr 141XO	Uterus 2	0.06	0.00
20	0= Negative			

In the analysis of matching samples, higher expression of Lng128 is detected in lung samples showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal pooled samples (Table 1).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows overexpression of Lng128 in 38 lung cancer tissues compared with their respective normal adjacent (lung samples #1, 2, 4, 5, 7, 9, 10, 12, 13, 14, 17, 19, 21, 23, 24, 25, 27, 28, 30, 31, 35 and 35,). There is overexpression in the cancer tissue for 55% of the lung matching samples tested (21 out of total of 38 lung matching samples).

Altogether, the high level of tissue specificity, plus the mRNA differential expression in the lung matching samples tested are believed to make Lng128 a good marker for diagnosing, monitoring, staging, imaging and treating 5 lung cancer.

**Northern Analysis**

Two transcripts ~ 2.2 kb and ~ 4.2 kb

**Primers Used for QPCR Expression Analysis**10 **Forward primer**

CTTGGTCTTCCTGCTCCTGAC (SEQ ID NO:21)

**Reverse primer**

AGGGCAGAGAGGAACAGCA (SEQ ID NO:22)

**Probe**

## 15 CCAGCGAGGAGCAGCAGGGATG (SEQ ID NO:23)

**Example 2****Sequence 2****Lng129****Gene ID 347842**

20 **Table 1.** The absolute numbers are relative levels of expression of Lng129 in 24 normal different tissues. All the values are compared to normal spleen (calibrator). These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from 25 different individuals.

Tissue	NORMAL
Adrenal Gland	0.00
Bladder	0.00
Brain	0.00
Cervix	0.02
Colon	0.00
Endometrium	0.03
Esophagus	0.00
Heart	0.00
Kidney	0.00
Liver	0.01
Lung	0.12

	Mammary Gland	0.00
	Muscle	0.00
	Ovary	0.04
	Pancreas	0.00
5	Prostate	0.01
	Rectum	0.00
	Small Intestine	0.00
	Spleen	1.00
	Stomach	0.00
10	Testis	0.01
	Thymus	0.03
	Trachea	0.06
	Uterus	0.06

0=negative

15 The relative levels of expression in Table 1 show that Lng129 mRNA expression is high compared with most other normal tissues analyzed.

15 The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from 20 different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

25 **Table 2.** The absolute numbers are relative levels of expression of Lng129 in 67 pairs of matching samples and 1 ovary normal and one ovary cancer sample. All the values 25 are compared to normal spleen (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING	NORMAL
30	Lng 60L	Adenocarcinoma	Lung 1	0.71	0.69	
	Lng 143L	Adenocarcinoma	Lung 2	0.01	0.00	
	Lng 60XL	Adenocarcinoma	Lung 3	0.00	0.01	
	Lng AC82	Adenocarcinoma	Lung 4	0.31	0.00	
35	Lng AC88	Adenocarcinoma	Lung 5	0.00	0.00	
	Lng AC66	Adenocarcinoma	Lung 6	0.80	0.07	
	Lng AC69	Adenocarcinoma	Lung 7	0.00	0.00	
	Lng AC11	Adenocarcinoma	Lung 8	0.48	0.06	
40	Lng AC32	Adenocarcinoma	Lung 9	0.39	0.00	
	Lng AC39	Adenocarcinoma	Lung 10	0.53	0.01	
	Lng AC94	Adenocarcinoma	Lung 11	0.05	0.03	
	Lng AC90	Adenocarcinoma	Lung 12	0.04	0.00	
	Lng 47XQ	Adenocarcinoma	Lung 13	0.12	0.00	
45	Lng 223L	Adenocarcinoma	Lung 14	0.04	0.00	
	Lng 528L	Adenocarcinoma	Lung 15	0.00	0.00	
	Lng BR26	Bronchio-alveolar	Lung 16	0.24	0.94	

	Lng BA641	Bronchogenic carcinoma	Lung 17	0.40	0.10
	Lng 315L	Squamous cell	Lung 18	0.03	0.12
	Lng SQ45	Squamous cell	Lung 19	0.00	0.00
	Lng SQ14	Squamous cell	Lung 20	0.02	0.11
5	Lng SQ9X	Squamous cell	Lung 21	0.00	0.00
	Lng SQ56	Squamous cell	Lung 22	0.43	0.12
	Lng SQ80	Squamous cell	Lung 23	0.00	0.00
	Lng SQ32	Squamous cell	Lung 24	0.06	0.00
	Lng SQ16	Squamous cell	Lung 25	0.01	0.00
10	Lng SQ79	Squamous cell	Lung 26	0.11	0.04
	Lng 90X	Squamous cell	Lung 27	0.00	0.00
	Lng BR94	Squamous cell	Lung 28	4.76	0.00
	Lng C20X	Squamous cell	Lung 29	0.00	0.00
	Lng SQ44	Squamous cell	Lung 30	0.04	0.00
15	Lng SQ43	Squamous cell	Lung 31	0.82	0.08
	Lng 77L	Large cell carcinoma	Lung 32	0.00	0.00
	Lng LC71	Large cell carcinoma	Lung 33	0.05	0.30
	Lng LC109	Large cell carcinoma	Lung 34	1.48	0.90
20	Lng LC80	Large cell carcinoma	Lung 35	1.09	0.00
	Lng 75XC	Metastatic from bone	Lung 36	0.00	0.00
	Lng MT71	Metastatic from renal	Lung 37	0.18	0.04
	Lng MT67	Metastatic from	Lung 38	0.55	0.04
	Bld 46XK		Bladder 1	0.02	0.00
	Bld TR14		Bladder 2	0.46	0.39
25	Cvx KS52		Cervix 1	0.26	0.03
	Cvx KS83		Cervix 2	0.00	0.00
	ClnAS45		Colon 1	0.00	0.00
	ClnRC01		Colon 2	0.01	0.02
	End		Endometrium	0.00	0.00
30	kid 107XD		Kidney 1	1.53	0.03
	Kid 109XD		Kidney 2	0.33	0.11
	Liv 175L		Liver 1	0.27	0.03
	Liv174L		Liver 2	0.01	0.01
	Mam		Mammary 1	0.02	0.01
35	Mam 497M		Mammary 2	0.00	0.00
	Ovr A082		Ovary 1	0.01	0.00
	Ovr 18GA		Ovary 2		0.01
	Ovr 180B		Ovary 3	0.00	
	Pan 77X		Pancreas 1	0.00	0.00
40	Pan		Pancreas 2	0.00	0.00
	Pro		Prostate 1	0.01	0.02
	Pro		Prostate 2	0.00	0.00
	Skn 248S		Skin 1	0.13	0.02
	SmInt		Small	0.02	0.01
45	SmInt H89		Small	0.00	0.00
	Sto		Stomach 1	0.15	0.01
	Sto 531S		Stomach 2	0.00	0.00
	Tst647T		Testis 1	0.00	0.00
	Tst 39X		Testis 2	0.30	0.02
50	Thr		Thyroid 1	0.04	0.03
	Thr 270T		Thyroid 2	0.11	0.00
	Utr135XO		Uterus 1	0.20	0.00

Utr 141XO	Uterus 2	0.00	0.01
0= Negative			

In the analysis of matching samples, higher expression  
5 of Lng129 is detected in lung samples showing a high degree  
of tissue specificity for lung tissue. These results  
confirm the tissue specificity results obtained with normal  
pooled samples (Table 1).

Furthermore, we compared the level of mRNA expression  
10 in cancer samples and the isogenic normal adjacent tissue  
from the same individual. This comparison provides an  
indication of specificity for the cancer stage (e.g. higher  
levels of mRNA expression in the cancer sample compared to  
the normal adjacent). Table 2 shows overexpression of  
15 Lng129 in 38 lung cancer tissues compared with their  
respective normal adjacent (lung samples #2, 4, 6, 8, 9,  
10, 11, 12, 13, 14, 17, 22, 24, 25, 26, 28, 30, 31, 33, 34,  
35, 37, and 38). There is overexpression in the cancer  
tissue for 61% of the lung matching samples tested (23 out  
20 of total of 38 lung matching samples).

Altogether, the high level of tissue specificity, plus  
the mRNA differential expression in the lung matching  
samples tested are believed to make Lng129 a good marker  
for diagnosing, monitoring, staging, imaging and treating  
25 lung cancer.

#### **Northern Analysis**

Two transcripts ~ 6.5 kb and ~ 9 kb

#### **DNA sequence for Lng129**

Sequence available from Incyte database.

#### **30 Primers Used for QPCR Expression Analysis**

##### **Forward primer**

GCCTGTTGGGAGATTAGATT (SEQ ID NO:24)

**Reverse primer**

GCCCAACAGAACAGACTAAAAA (SEQ ID NO:25)

**Probe**

AGGTTATTAGGTTATTATCTCTCTCCTGATTTTCC (SEQ ID NO:26)

5 **Example 3****Sequence 3****Lng112****Gene ID 983439**

Table 1. The absolute numbers are relative levels of expression of Lng112 in 12 normal different tissues. These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

	Tissue	NORMAL
15	Brain	0
	Heart	0
	Kidney	0
	Liver	0
20	Lung	1.0
	Mammary	0
	Muscle	0
	Prostate	0
25	SmInt	0
	Testis	0
	Thymus	0
	Uterus	0

0=negative

The relative levels of expression in Table 1 show that Lng112 mRNA expression is only detectable in lung compared with other normal tissues analyzed

The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

Table 2. The absolute numbers are relative levels of expression of Lng112 in 49 pairs of matching samples. All the values are compared to normal lung (calibrator). A

matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING NORMAL ADJACENT
5	Lung 143L	Adenocarcinoma	Lung 1	0.00	0.00
	Lung 60L	Adenocarcinoma	Lung 2	13.00	9.00
	Lng AC82	Adenocarcinoma	Lung 3	0.00	5.00
10	Lng 60XL	Adenocarcinoma	Lung 4	0.00	19.00
	Lng AC66	Adenocarcinoma	Lung 5	0.00	59.00
	Lng AC69	Adenocarcinoma	Lung 6	9.00	0.00
	Lng AC88	Adenocarcinoma	Lung 7	16.00	35.00
15	Lng AC11	Adenocarcinoma	Lung 8	0.00	36.00
	Lng AC32	Adenocarcinoma	Lung 9	0.00	41.00
	Lng AC39	Adenocarcinoma	Lung 10	0.00	5.00
	Lng AC94	Adenocarcinoma	Lung 11	0.00	0.00
	BA641	Bronchio-alveolar carcinoma	Lung 12	7.00	14.00
	Lng SQ32	Squamous cell carcinoma	Lung 13	0.00	228.00
20	Lng SQ45	Squamous cell carcinoma	Lung 14	368.00	2.00
	Lng SQ56	Squamous cell carcinoma	Lung 15	1.00	53.00
	Lng SQ9X	Squamous cell carcinoma	Lung 16	0.00	2.00
	Lng SQ14	Squamous cell carcinoma	Lung 17	0.00	21.00
	Lng SQ16	Squamous cell carcinoma	Lung 18	0.01	1.00
25	Lng SQ80	Squamous cell carcinoma	Lung 19	6.00	7.00
	Lng C20X	Squamous cell carcinoma	Lung 20	0.00	0.00
	Lng 47XQ	Squamous cell carcinoma	Lung 21	1.00	3.00
	Lng SQ44	Squamous cell carcinoma	Lung 22	0.00	0.00
	Lng SQ79	Squamous cell carcinoma	Lung 23	0.00	0.00
30	Lng 90X	Squamous cell carcinoma	Lung 24	0.00	4.00
	Lng BR94	Squamous cell carcinoma	Lung 25	0.00	0.00
	Lng LC71	Large cell carcinoma	Lung 26	178.00	4.00
	Lng LC80	Large cell carcinoma	Lung 27	0.00	0.00
35	Lng LC109	Large cell carcinoma	Lung 28	1.00	96.00
	Lung 77L	Large cell carcinoma	Lung 29	0.00	0.00
	Lng 75XC	Metastatic from bone cancer	Lung 30	0.00	22.00
	Lng MT67	Metastatic from renal cell cancer	Lung 31	1.00	86.00
	Lng MT71	Metastatic from melanoma	Lung 32	0.00	14.00
40	Bld 32XK		Bladder 1	0.00	0.00
	Cln AS45		Colon 1	0.00	0.00

	Cvx	KS52	Cervix 1	0.00	0.00
	End	28XA	Endometrium 1	0.00	0.00
	Kid		Kidney 1	0.00	0.00
	106XD				
5	Liv	94XA	Liver 1	0.00	0.00
	Mam	A06X	Mammary 1	0.00	0.00
	Ovr	103X	Ovary 1	0.00	0.00
	Pan	71XL	Pancreas 1	0.00	0.00
	Pan	77X	Pancreas 2	0.00	0.00
10	Pro	20XB	Prostate 1	0.00	0.00
	Skn	287S	Skin 1	0.00	0.00
	SmInt		Sm. Int. 1	0.00	0.00
	H89				
15	Sto	531S	Stomach 1	0.00	0.00
	Thr	143N	Thyriod 1	0.00	0.00
	Tst	39X	Testis 1	0.00	0.00
	Utr		Uterus 1	16.00	0.00
	135XO				

0= Negative

20 In the analysis of matching samples, except 1 uterus cancer sample the only detection was in lung samples showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal pooled samples (Table 1).

25 Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher levels of mRNA expression in the cancer sample compared to 30 the normal adjacent). Table 2 shows overexpression of Lng112 in 4 lung cancer tissues compared with their respective normal adjacent tissue in 32 cancer matching pairs (lung samples # 2, 6, 14, and 26). There is 35 overexpression in the cancer tissue for 12.5% of the lung matching samples tested (total of 32 lung matching samples).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. The result shows that Lng112 is expressed differentially in all 32 lung cancer tissues 5 tested compared with their respective normal adjacent.

Altogether, the high level of tissue specificity, plus the mRNA differential expression in the lung matching samples tested are believed to make Lng112 a good marker for diagnosing, monitoring, staging, imaging and treating 10 lung cancer.

**Primers Used for QPCR Expression Analysis**

**Forward primer**

TGGTGGCGTTCCCTCGTC (SEQ ID NO:27)

**Reverse primer**

15 CAGAGCCCTTCGTACTGGAACAC (SEQ ID NO:28)

**Probe**

TCGTACAGGTCCCTGGGTGCTCCACA (SEQ ID NO:29)

**Example 4**

**Sequence 4**

20 **Lng114**

Gene ID 236582

**Table 1.** The absolute numbers are relative levels of expression of Lng114 in 12 normal different tissues. All the values are compared to normal testis (calibrator). 25 These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

Tissue	NORMAL
Brain	0.09
Heart	0.00
Kidney	0.26
Liver	0.00
Lung	602.58
Mammary	0.35

5	Muscle	0.00
	Prostate	0.00
	SmInt	0.05
	Testis	1.00
	Thymus	0.00
	Uterus	1.27

0=negative

The relative levels of expression in Table 1 show that Lng114 mRNA expression is highest in lung (602.58) compared 10 with other normal tissues analyzed.

The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue 15 samples of a single individual in Table 2.

**Table 2.** The absolute numbers are relative levels of expression of Lng114 in 78 pairs of matching samples, 1 normal ovary and 2 blood samples. All the values are compared to normal testis (calibrator). A matching pair is 20 formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

25	Sample ID	Cancer Type	Tissue	CANCER	NORMAL NORMAL ADJACENT
Lng 60L	Adenocarcinoma	Lung 1	121.52	66.72	
Lng 143L	Adenocarcinoma	Lung 2	360.79	25.99	
Lng 60XL	Adenocarcinoma	Lung 3	79.81	648.73	
Lng AC82	Adenocarcinoma	Lung 4	37.53	102.18	
Lng AC88	Adenocarcinoma	Lung 5	530.06	992.55	
Lng AC66	Adenocarcinoma	Lung 6	76.68	257.93	
Lng AC69	Adenocarcinoma	Lung 7	25.46	8.51	
Lng AC11	Adenocarcinoma	Lung 8	54.19	852.17	
Lng AC32	Adenocarcinoma	Lung 9	157.59	193.34	
Lng AC94	Adenocarcinoma	Lung 10	2272.40	112.99	
Lng AC90	Adenocarcinoma	Lung 11	141.53	38.85	
Lng AC39	Adenocarcinoma	Lung 12	198.09	8.40	
Lng 223L	Adenocarcinoma	Lung 13	10.63	31.89	
Lng 528L	Adenocarcinoma	Lung 14	210.84	274.37	
Lng BR26	Bronchogenic carcinoma	Lung 15	0.00	169.48	

	Lng BA641	Bronchio-alveolar carcinoma	Lung 16	316.27	73.77
	Lng 315L	Squamous cell carcinoma	Lung 17	0.00	469.51
5	Lng SQ14	Squamous cell carcinoma	Lung 18	1016.93	15.83
	Lng SQ56	Squamous cell carcinoma	Lung 19	0.78	526.39
	Lng SQ9X	Squamous cell carcinoma	Lung 20	52.89	64.89
	Lng SQ80	Squamous cell carcinoma	Lung 21	60.34	962.07
	Lng SQ45	Squamous cell carcinoma	Lung 22	97.01	357.05
10	Lng SQ16	Squamous cell carcinoma	Lung 23	92.41	1833.01
	Lng SQ32	Squamous cell carcinoma	Lung 24	23.75	31.02
	Lng SQ79	Squamous cell carcinoma	Lung 25	20.89	142.52
	Lng 47XQ	Squamous cell carcinoma	Lung 26	42.52	135.77
	Lng BR94	Squamous cell carcinoma	Lung 27	211.50	157.78
	Lng 90X	Squamous cell carcinoma	Lung 28	80.73	12.21
15	Lng C20X	Squamous cell carcinoma	Lung 29	2.99	15.24
	Lng SQ44	Squamous cell carcinoma	Lung 30	94.03	0.00
	Lng SQ43	Squamous cell carcinoma	Lung 31	27.19	38.85
	Lng LC71	Large cell carcinoma	Lung 32	1217.75	2040.91
20	Lng LC109	Large cell carcinoma	Lung 33	160.42	4576.44
	Lng LC80	Large cell carcinoma	Lung 34	955.43	400.32
	Lng 77L	Large cell carcinoma	Lung 35	18.44	78.52
	Lng 75XC	Metastatic from bone cancer	Lung 36	229.13	398.93
25	Lng MT67	Metastatic from renal cell cancer	Lung 37	69.07	1514.89
	Lng MT71	Metastatic from melanoma	Lung 38	42.37	1393.99
30	Bld 46XK		Bladder 1	0.00	0.00
	Bld 66X		Bladder 2	0.00	0.00
	Blo B5		Blood 1		0.00
	Blo B6		Blood 2		0.00
35	Cln AS43		Colon 1	1.22	0.00
	Cln AS45		Colon 2	0.00	0.00
	Cln AS46		Colon 3	2.08	0.00
	Cln SG67		Colon 4	1.49	1.39
	Cvx KS52		Cervix 1	2.39	11.47
40	Cvx KS83		Cervix 2	1.22	4.55
	Endo 28XA		Endometrium 1	108.38	2.86
	Endo 68X		Endometrium 2	3.73	12.64
	Kid10XD		Kidney 1	39.40	0.00

DO NOT CONFIDE

	Kid 109XD	Kidney 2	1.91	8.46
	Kid 107XD	Kidney 3	1.48	4.61
5	Liv 15XA	Liver 1	0.03	0.07
	Liv 201L	Liver 2	0.00	0.00
	Liv 174L	Liver 3	0.00	0.00
	Mam 162X	Mammary 1	0.78	0.28
10	Mam 173M	Mammary 2	1.00	0.00
	Mam 220	Mammary 3	2.02	0.30
	Ovr 18GA	Ovary 1		0.00
	Ovr AO84	Ovary 2	0.00	0.00
	Pro	Prostate 1	0.86	1.38
15	101XB			
	Pro	Prostate 2	0.00	0.23
	109XB			
	Pro125XB	Prostate 3	0.00	0.08
	Pan 77X	Pancreas 1	0.00	0.00
20	Skn 39A	Skin 1	0.20	0.00
	Skn 39AB	Skin 2	0.00	0.00
	Skn 248S	Skin 3	0.00	0.00
	Smint 21XA	Sm. Int. 1	0.110	0.00
	Smint			
25	H89	Sm. Int. 2	0.00	0.00
	Sto 264S	Stomach 1	1.04	5.98
	Sto 288S	Stomach 2	5.10	0.00
	Sto 115S	Stomach 3	5.03	0.75
30	Thr 143N	Thyroid 1	0.00	0.94
	Thr 145T	Thyroid 2	1.89	2.50
	Thr 939T	Thyriod 3	1.52	0.00
	Tst 647T	Testis 1	10.20	0.00
35	Tst 39X	Testis 2	8.20	0.00
	Tst 663T	Testis 3	5.09	0.00
	Utr 141XO	Uterus 1	13.18	5.65
40	Utr 135XO	Uterus 2	1.47	1.36

0= Negative

In the analysis of matching samples, higher expression of lng114 is detected in lung samples showing a high degree of tissue specificity for lung tissue. These results 45 confirm the tissue specificity results obtained with normal pooled samples (Table 1).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an 50 indication of specificity for the cancer stage (e.g. higher levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows overexpression of

Lng114 in 11 lung cancer tissues compared with their respective normal adjacent tissue in 35 cancer matching pairs (lung samples #1, 2, 7, 10, 11, 12, 15, 16, 26, 28, and 32). There is overexpression in the cancer tissue for 5 31% of the lung matching samples tested (total of 35 primary cancer lung matching samples).

Altogether, the high level of tissue specificity, plus the mRNA overexpression in 31% of the lung matching samples tested are believed to make Lng114 a good marker for 10 diagnosing, monitoring, staging, imaging and treating lung cancer.

#### Primers Used for QPCR Expression Analysis

##### **Forward primer**

CTTGGCAGCTCACATGGAAC (SEQ ID NO:30)

15 **Reverse primer**

CTGGGGTGTCTCTGTCACTCTC (SEQ ID NO:31)

##### **Probe**

CCATGAAGTCCCACCCCTTTCTCTG (SEQ ID NO:32)

#### **Example 5**

20 **Sequence 5**

##### **Lng118**

Gene ID 210995

Table 1. The absolute numbers are relative levels of expression of Lng118 in 24 normal different tissues. These 25 RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

Tissue	NORMAL
Adrenal Gland	0
Bladder	0
Brain	0.010
Cervix	0.010
Colon	0
Endometrium	0.010

	Esophagus	0
	Heart	0
	Kidney	0.010
	Liver	0
5	Lung	1.000
	Mammary Gland	0.010
	Muscle	0.0032
	Ovary	0.005
	Pancreas	0.005
10	Prostate	0.002
	Rectum	0.004
	Small Intestine	0
	Spleen	0
	Stomach	0.015
15	Testis	0.033
	Thymus	0.001
	Trachea	0.007
	Uterus	0.005

0=negative

20 The relative levels of expression in Table 1 show that Lng118 mRNA expression is high in lung compared with other normal tissues analyzed.

25 The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

30 **Table 2.** The absolute numbers are relative levels of expression of Lng118 in 36 pairs of matching samples. All the values are compared to normal lung (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING
35	Lng60L	Adenocarcinoma	Lung 1	0.1	0.04
	LngAC66	Adenocarcinoma	Lung 2	0	0.27
	LngAC69	Adenocarcinoma	Lung 3	0	0.11
	Lng AC88	Adenocarcinoma	Lung 4	0.05	0.13
	Lng 60XL	Adenocarcinoma	Lung 5	0	0.08
40	LngAC94	Adenocarcinoma	Lung 6	0	0

5	LngAC11	Adenocarcinoma	Lung 7	0	0.02
	LngAC32	Adenocarcinoma	Lung 8	0	0.05
	Lng 47XQ	Adenocarcinoma	Lung 9	0	0
	Lng223L	Adenocarcinoma	Lung 10	0	0.01
10	Lng BR26	Bronchio-alveolar	Lung 11	0	0
	LngSQ45	Squamous cell	Lung 12	0.54	0
	LngSQ16	Squamous cell	Lung 13	0	0
	LngSQ79	Squamous cell	Lung 14	0	0
15	Lng LC71	Large cell carcinoma	Lung 15	1.23	0.06
	Lng LC109	Large cell carcinoma	Lung 16	0	0.06
	Lng 75XC	Metastatic from	Lung 17	0	0
	BldTR17		Bladder 1	0	0
	Cvx KS52		Cervix 1	0	0
	ClnSG45		Colon 1	0	0
20	End 10479		Colon 2	0	0
	Kid 106XD		Endometrium 1	0	0
	Kid 5XD		Kidney 1	0	0.01
	Liv 187L		Kidney 2	0	0
	Liv175L		Liver 1	0	0
25	Mam S967		Liver 2	0	0
	Ovr A084		Mammary 1	0	0
	Pan 71XL		Ovary 1	0.14	0
	Pro 20XB		Pancreas 1	0	0
	Pro 326		Prostate 1	0.02	0
30	SmInt H89		Prostate 2	0	0
	Sto 531S		Small	0	0
	Tst 39X		Stomach 1	0	0
	Thr 270T		Testis 1	0	0
	Thr 644T		Thyroid 1	0.02	0.01
	0= Negative				

In the analysis of matching samples, higher expression of lng118 is detected in lung samples showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal pooled samples (Table 1).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows differential expression of Lng118 in 17 lung cancer tissues compared with their respective normal adjacent tissue.

Altogether, the high level of tissue specificity, plus the mRNA differential expression in the lung matching samples tested are believed to make Lng118 a good marker for diagnosing, monitoring, staging, imaging and treating 5 lung cancer.

**DNA sequence for Lng118**

Sequence available from Incyte database.

**Primers Used for QPCR Expression Analysis****Forward primer**

10 TGCAGCAGAAAGGGGAGAG (SEQ ID NO:33)

**Reverse primer**

TCCCCATTGCCCTCAAGT (SEQ ID NO:34)

**Probe**

CGTGGGCACTCACCTCGGCAC (SEQ ID NO:35)

15 **Example 6**

**Sequence 6****Lng121**

Gene ID 208994

Table 1. The absolute numbers are relative levels of 20 expression of Lng121 in 24 normal different tissues. All the values are compared to normal trachea (calibrator). These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

25

Tissue	NORMAL
Adrenal Gland	0.01
Bladder	0.00
Brain	0.55
Cervix	0.09
Colon	0.02
Endometrium	1.74
Esophagus	0.08
Heart	0.00
Kidney	0.04

30

	Liver	0.00
	Lung	117.38
	Mammary Gland	0.47
	Muscle	0.36
5	Ovary	0.41
	Pancreas	0.10
	Prostate	0.93
	Rectum	0.05
10	Small Intestine	0.09
	Spleen	1.72
	Stomach	0.12
	Testis	3.24
	Thymus	2.06
	Trachea	1.00
15	Uterus	0.12

0=negative

The relative levels of expression in Table 1 show that Lng121 mRNA expression is high in lung compared with most other normal tissues analyzed.

20 The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

25 **Table 2.** The absolute numbers are relative levels of expression of Lng121 in 20 pairs of matching samples. All the values are compared to normal trachea (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent

30 sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING NORMAL ADJACENT
	Lng 60L	Adenocarcinoma	Lung 1	9.95	16.62
	Lng 143L	Adenocarcinoma	Lung 2	1.07	5.41
	Lng 60XL	Adenocarcinoma	Lung 3	12.77	5.08
35	Lng AC88	Adenocarcinoma	Lung 4	5.06	31.89
	Lng AC66	Adenocarcinoma	Lung 5	3.85	22.32
	Lng AC32	Adenocarcinoma	Lung 6	8.46	87.12
	Lng 223L	Adenocarcinoma	Lung 7	1.87	4.10
	Lng SQ14	Squamous cell carcinoma	Lung 8	2.91	33.72
40	Lng C20X	Squamous cell carcinoma	Lung 9	0.08	0.29

5	Lng 77L	Large cell carcinoma	Lung 10	8.13	16.35
	Lng LC71	Large cell carcinoma	Lung 11	47.84	3.69
	Lng 75XC	Metastatic from melanoma	Lung 12	3.49	15.67
	Cln AS43		Colon 1	1.22	0.17
	Endo 12XA		Endometrium	2.38	0.29
	Kid 107XD		1		
	Liv 187L		Kidney 1	0.44	0.17
	Mam 19DN		Liver 1	0.03	1.06
	Ovr A084		Mammary 1	1.41	0.58
10	Pro 109XB		Ovary 1	0.76	0.28
	Tst 647T		Prostate 1	0.19	0.27
			Testis 1	2.92	1.64

0= Negative

In the analysis of matching samples, higher expression of lng121 is detected in lung samples showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal pooled samples (Table 1).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows differential expression of Lng121 in 12 lung cancer tissues compared with their respective normal adjacent tissue.

Altogether, the high level of tissue specificity, plus the mRNA differential expression in the lung matching samples tested are believed to make Lng121 a good marker for diagnosing, monitoring, staging, imaging and treating lung cancer.

#### DNA sequence for Lng121

Sequence available from Incyte database.

#### Primers Used for QPCR Expression Analysis

##### Forward primer

35 CAGGCTCATTTATTGTGGTCAT (SEQ ID NO:36)

##### Reverse primer

CCCACACTGATTTAGGCACATAG (SEQ ID NO:37)

**Probe**

TTTGAAGGAGGGCAGGAAAACTATGTAAG (SEQ ID NO:38)

**Example 7****Sequence 7**5 **Lng124**

Gene ID 1066498

**Table 1.** The absolute numbers are relative levels of expression of Lng124 in 24 normal different tissues. All the values are compared to normal lung (calibrator). These 10 RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

	Tissue	NORMAL
15	Adrenal Gland	0
	Bladder	0
	Brain	0
	Cervix	0
	Colon	0
	Endometrium	0
20	Esophagus	0
	Heart	0
	Kidney	0
	Liver	0
	Lung	1.00
25	Mammary Gland	0
	Muscle	0
	Ovary	0
	Pancreas	0
	Prostate	0
30	Rectum	0
	Small Intestine	0
	Spleen	0
	Stomach	0
	Testis	0
35	Thymus	0
	Trachea	0
	Uterus	0

0=negative

The relative levels of expression in Table 1 show that 40 Lng124 mRNA expression is only detectable in lung compared with most other normal tissues analyzed.

The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue

5 samples of a single individual in Table 2.

**Table 2.** The absolute numbers are relative levels of expression of Lng124 in 40 pairs of matching samples. All the values are compared to normal lung (calibrator). A matching pair is formed by mRNA from the cancer sample for

10 a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING NORMAL ADJACENT
15	Lng60L	Adenocarcinoma	Lung 1	0.64	0.62
	Lng 143L	Adenocarcinoma	Lung 2	0.03	0.26
	LngAC66	Adenocarcinoma	Lung 3	0.04	0.51
	Lng 60XL	Adenocarcinoma	Lung 4	0.04	0.36
	Lng AC88	Adenocarcinoma	Lung 5	0.11	1.05
20	LngAC11	Adenocarcinoma	Lung 6	0.12	2.48
	LngAC32	Adenocarcinoma	Lung 7	0.22	0.64
	Lng 47XQ	Adenocarcinoma	Lung 8	0.19	0.13
	Lng AC39	Adenocarcinoma	Lung 9	0.16	0.7
	Lng AC90	Adenocarcinoma	Lung 10	0.1	0.12
25	Lng223L	Adenocarcinoma	Lung 11	0.05	0.17
	Lng SQ14	Squamous cell carcinoma	Lung 12	0	0.69
	Lng SQ9X	Squamous cell carcinoma	Lung 13	0.15	0.12
	LngSQ16	Squamous cell carcinoma	Lung 14	0.12	0.24
	LngSQ79	Squamous cell carcinoma	Lung 15	0.1	0.42
30	Lng SQ43	Squamous cell carcinoma	Lung 16	0.14	0.17
	Lng BR94	Squamous cell carcinoma	Lung 17	0.01	0.03
	Lng C20X	Squamous cell carcinoma	Lung 18	0	0.02
	Lng LC109	Large cell carcinoma	Lung 19	0.06	0.74
	Bld 66X		Bladder 1	0	0
35	Cvx NK23		Cervix 1	0	0
	Cvx NK24		Cervix 2	0	0
	ClnAS45		Colon 2	0	0
	Cln RC24		Colon 3	0	0
	End 8911		Endometrium 1	0	0
40	Kid 6XD		Kidney 1	0	0
	Kid 710K		Kidney 2	0	0
	Liv 94XA		Liver 1	0	0
	Mam 173M		Mammary 2	0	0
	Mam S123		Mammary 3	0	0

	Ovr A082	Ovary 1	0	0
	Ovr C179	Ovary 2	0	0
	Ovr 130X	Ovary 3	0	
	Pan 92X	Pancreas 1	0	0
5	Pro 34B	Prostate 1	0	0
	Sto 531S	Stomach 1	0	0
	Sto AC93	Stomach 2	0	0
	Sto 288S	Stomach 3	0	0
	Sto TA73	Stomach 4	0	0
10	Sto 288S	Stomach 5	0	0
	Sto 531S	Stomach 6	0	0
	Skn 287S	Skin 1	0	0
	Thr692T	Thyroid 1	0.02	0

0= Negative

15 In the analysis of matching samples, expression of Lng124 is only detected in lung samples (except 1 thyroid cancer sample) showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal pooled samples

20 (Table 1).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher 25 levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows differential expression of Lng124 in all of the 19 lung cancer tissues compared with their respective normal adjacent tissue.

Altogether, the high level of tissue specificity, plus 30 the mRNA differential expression in all of the lung matching samples tested are believed to make Lng124 a good marker for diagnosing, monitoring, staging, imaging and treating lung cancer.

#### Primers Used for QPCR Expression Analysis

35 **Forward primer**

AGGGAGAGGAGCTATGGACGT (SEQ ID NO:39)

**Reverse primer**

TTTGAGGCAAGACTCCATCTC (SEQ ID NO:40)

**Probe**

CTGCCAAGGGAGAGAGTGAGGTAGGC (SEQ ID NO:41)

**Example 8****Sequence 8**5 **Lng126**

Gene ID 287016

**Table 1.** The absolute numbers are relative levels of expression of Lng126 in 24 normal different tissues. All the values are compared to normal thymus (calibrator).

10 These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

	Tissue	NORMAL
15	Adrenal Gland	11.92
	Bladder	0.0
	Brain	0.21
	Cervix	0.7
	Colon	0.06
20	Endometrium	6.36
	Esophagus	0.04
	Heart	0.06
	Kidney	1.11
	Liver	7.94
	Lung	6.2
25	Mammary Gland	7.46
	Muscle	0.78
	Ovary	38.32
	Pancreas	2.69
	Prostate	5.21
30	Rectum	2.72
	Small	0.6
	Spleen	0.16
	Stomach	0.93
	Testis	3.2
35	Thymus	1.00
	Trachea	4.61
	Uterus	3.90

0=negative

The relative levels of expression in Table 1 show that

40 Lng126 mRNA expression is relatively high in lung, except

adrenal gland and ovary, compared with other normal tissues analyzed.

The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from 5 different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

**Table 2.** The absolute numbers are relative levels of expression of Lng126 in 20 pairs of matching samples. All 10 the values are compared to normal thymus (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING NORMAL ADJACENT
15	Lng 60L	Adenocarcinoma	Lung 1	0.06	0.03
	Lng 143L	Adenocarcinoma	Lung 2	7.34	0.45
	LngAC66	Adenocarcinoma	Lung 3	0.10	0.07
	LngAC69	Adenocarcinoma	Lung 4	0.33	0.04
20	LngAC11	Adenocarcinoma	Lung 5	0.72	0.30
	LngAC32	Adenocarcinoma	Lung 6	0.14	0.10
	LngAC94	Adenocarcinoma	Lung 7	0.11	0.01
	Lng223L	Adenocarcinoma	Lung 8	0.01	0.01
	LngSQ45	Squamous cell carcinoma	Lung 9	0.43	0.16
	Lng SQ14	Squamous cell carcinoma	Lung 10	11.35	2.61
25	LngSQ16	Squamous cell carcinoma	Lung 11	0.09	0.01
	LngSQ79	Squamous cell carcinoma	Lung 12	10.78	0.14
	Lng C20X	Squamous cell carcinoma	Lung 13	0.26	0.00
30	Lng 77L	Large cell carcinoma	Lung 14	1.32	7.14
	Bld 66X		Bladder 1	4.92	43.56
	ClnAS45		Colon 1	1.26	1.28
	Mam 19DN		Mammary 1	14.62	0.48
	Mam 220		Mammary 2	0.33	0.61
	Mam S854		Mammary 3	0.66	1.04

0= Negative

In the analysis of matching samples, higher expression of lng126 is detected in lung samples showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal 5 pooled samples (Table 1).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher 10 levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows differential expression of Lng126 in 14 lung cancer tissues compared with their respective normal adjacent tissue.

Altogether, the high level of tissue specificity, plus 15 the mRNA differential expression in the lung matching samples tested are believed to make Lng126 a good marker for diagnosing, monitoring, staging, imaging and treating lung cancer.

#### **DNA sequence for Lng126**

20 Sequence available from Incyte database.

#### **Primers Used for QPCR Expression Analysis**

##### **Forward primer**

TGGGGACAATATGGACCTCA (SEQ ID NO:42)

##### **25 Reverse primer**

GGCGAGTGTCTATGATGAAACCT (SEQ ID NO:43)

##### **Probe**

CAGGATCTGTGAGGATTCATTTGGATACAT (SEQ ID NO:44)

#### **Example 9**

30 **Sequence 9**

**Lng136**

Gene ID 10717

ddx lung code SQLng001

Table 1. The absolute numbers are relative levels of expression of Lng136 in 24 normal different tissues. All the values are compared to normal spleen (calibrator). These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

	Tissue	NORMAL
10	Adrenal Gland	0.34
	Bladder	0.03
	Brain	0.66
15	Cervix	0.12
	Colon	0.00
	Endometrium	0.08
	Esophagus	0.05
20	Heart	0.02
	Kidney	0.01
	Liver	0.00
	Lung	8.54
25	Mammary	1.32
	Muscle	0.00
	Ovary	0.07
30	Pancreas	0.86
	Prostate	0.15
	Rectum	0.02
	Small Int.	0.05
	Spleen	1.0
35	Stomach	0.77
	Testis	1.22
	Thymus	0.19
	Trachea	0.16
	Uterus	0.03

0=negative

The relative levels of expression in Table 1 show that Lng136 mRNA expression is high in lung compared with most other normal tissues analyzed.

The absolute numbers in Table 1 were obtained 40 analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the

absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

**Table 2.** The absolute numbers are relative levels of expression of Lng136 in 60 pairs of matching samples. All 5 the values are compared to normal spleen (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING NORMAL ADJACENT
10	Lng60L	Adenocarcinoma	Lung 1	5.92	4.77
	Lng 143L	Adenocarcinoma	Lung 2	2.20	2.58
	Lng AC82	Adenocarcinoma	Lung 3	2.62	6.68
	LngAC66	Adenocarcinoma	Lung 4	1.36	3.61
	Lng 60XL	Adenocarcinoma	Lung 5	0.66	4
15	Lng AC88	Adenocarcinoma	Lung 6	3.61	15.94
	LngAC69	Adenocarcinoma	Lung 7	7.49	7.19
	LngAC11	Adenocarcinoma	Lung 8	1.31	31.45
	LngAC32	Adenocarcinoma	Lung 9	7.41	9.51
20	Lng AC90	Adenocarcinoma	Lung 10	16.34	5.26
	Lng223L	Adenocarcinoma	Lung 11	2.41	1.49
	LngAC94	Adenocarcinoma	Lung 12	1.47	2.09
	Lng BR26	Bronchio-alveolar carcinoma	Lung 13	0	6.5
	LngSQ45	Squamous cell carcinoma	Lung 14	19.97	2.18
	Lng SQ14	Squamous cell carcinoma	Lung 15	1.76	13.04
25	Lng SQ56	Squamous cell carcinoma	Lung 16	2.18	12.73
	LngSQ16	Squamous cell carcinoma	Lung 17	0.54	5.30
	Lng SQ32	Squamous cell carcinoma	Lung 18	3.31	14.17
	Lng AC39	Squamous cell carcinoma	Lung 19	3.43	15.08
	Lng 47XQ	Squamous cell carcinoma	Lung 20	0.74	5.3
30	LngSQ79	Squamous cell carcinoma	Lung 21	2.53	8.49
	Lng C20X	Squamous cell carcinoma	Lung 22	0.07	0.22
	Lng SQ44	Squamous cell carcinoma	Lung 23	1.48	3.59
	Lng SQ43	Squamous cell carcinoma	Lung 24	1.45	0.91
35	Lng LC71	Large cell carcinoma	Lung 25	11.79	10.67
	Lng 77L	Large cell carcinoma	Lung 26	9.25	3.11
	Lng LC109	Large cell carcinoma	Lung 27	6.87	36.89
	Lng MT67	Metastatic from renal cell cancer	Lung 28	2.93	5.01
	Lng MT71	Metastatic from melanoma	Lung 29	0.19	1.23
	BldTR14		Bladder 1	0.25	0.94

	Cvx NK24	Cervix 1	0.46	0.14
	Cvx KS52	Cervix2	0.07	0.03
	ClnAS43	Colon1	0.02	0.07
	ClnAS45	Colon2	0.10	0.06
5	ClnAS46	Colon3	0.16	0.13
	ClnAS67	Colon4	0.04	0.14
	ClnAS89	Colon5	0.10	0.36
	End 8911	Endometrium	0.06	0.18
	1			
	End 28XA	Endometrium	0.23	0.12
	2			
10	Kid 5XD	Kidney1	0.01	1.41
	Kid 109XD	Kidney2	0.47	0.39
	Liv15XA	Liver1	0.11	0.03
	Liv 174L	Liver 2	0.01	0.01
15	Mam S123	Mammary 1	0.19	0.17
	Mam 162X	Mammary 2	0.15	0.17
	Ovr C179	Ovary 1	0	0
	Ovr 130X	Ovary 2	0.58	0
	Pan 71 XL	Pancreas 1	0.07	0
	Pan 92X	Pancreas 2	6.94	1.62
20	Pro 326	Prostate 1	0.04	0.12
	Pro 109XB	Prostate 2	0.01	0.01
	Skn 278S	Skin 1	0.04	0.1
	SmInt H89	Small	0.13	0.04
	SmInt	intestine 1		
25	21XA	Small	0.18	0
	Sto TA73	intestine 2		
	Sto 758S	Stomach 1	2.9	4.18
	Tst647T	Stomach 2	0.77	1.53
	Tst 39X	Testis 1	0.58	0.38
30	Thr 270T	Testis 2	0.50	1.02
	Utr135XO	Thyroid 1	0.03	0.02
	0= Negative	Uterus 1	0.2	0.48

In the analysis of matching samples, higher expression of lng136 is detected in lung samples showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal pooled samples (Table 1).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows differential expression of Lng136 in 29 lung cancer tissues compared with their respective normal adjacent tissue.

Altogether, the high level of tissue specificity, plus the mRNA differential expression in the lung matching

samples tested are believed to make Lng136 a good marker for diagnosing, monitoring, staging, imaging and treating lung cancer

**DNA sequence for Lng136**

5 Sequence available from Incyte database.

**Primers Used for QPCR Expression Analysis**

**Forward primer**

CTCCGTGGCTCGTGCTT (SEQ ID NO:45)

**Reverse primer**

10 CGCTTTCTTTGCCCTCTTGT (SEQ ID NO:46)

**Example 10**

**Sequence 10**

**Lng143**

15 Gene ID 24945

ddx lung code SQLng006

**Table 1.** The absolute numbers are relative levels of expression of Lng143 in 24 normal different tissues. All the values are compared to normal pancreas (calibrator).

20 These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

Tissue	NORMAL
Adrenal Gland	0.83
Bladder	0.04
Brain	1.11
Cervix	0.20
Colon	0.01
Endometrium	2.49
Esophagus	0.01
Heart	0.09
Kidney	0.34
Liver	0.23
Lung	6.15

	Mammary Gland	2.34
	Muscle	0.44
	Ovary	4.20
	Pancreas	1.00
5	Prostate	6.34
	Rectum	1.14
	Small Intestine	0.16
	Spleen	6.63
10	Stomach	1.13
	Testis	3.12
	Thymus	7.39
	Trachea	2.77
	Uterus	6.04

0=negative

15 The relative levels of expression in Table 1 show that Lng143 mRNA expression is much higher in lung compared with most other normal tissues analyzed.

15 The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from 20 different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

25 **Table 2.** The absolute numbers are relative levels of expression of Lng143 in 78 pairs of matching samples, 2 blood samples and 2 normal ovary samples. All the values 30 are compared to normal pancreas (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING NORMAL L	
					NORMAL	ADJACENT
	Lng 60L	Adenocarcinoma	Lung 1	0.54	0.41	
	Lng 143	Adenocarcinoma	Lung 2	0.41	0.08	
	Lng 60XL	Adenocarcinoma	Lung 3	1.09	0.86	
	Lng AC82	Adenocarcinoma	Lung 4	3.25	0.09	
35	Lng AC88	Adenocarcinoma	Lung 5	3.99	0.93	
	Lng AC66	Adenocarcinoma	Lung 6	0.99	0.42	
	Lng AC69	Adenocarcinoma	Lung 7	2.36	0.50	
	Lng AC11	Adenocarcinoma	Lung 8	2.67	1.80	
	Lng AC32	Adenocarcinoma	Lung 9	3.02	0.43	
40	Lng AC39	Adenocarcinoma	Lung 10	9.35	0.13	
	Lng AC94	Adenocarcinoma	Lung 11	0.58	0.26	
	Lng AC90	Adenocarcinoma	Lung 12	3.85	0.01	

	Lng 223L	Adenocarcinoma	Lung 13	0.32	0.02
	Lng 528L	Adenocarcinoma	Lung 14	10.52	3.77
	Lng BR26	Bronchogenic carcinoma	Lung 15	8.40	0.28
	Lng BA641	Bronchio-alveolar carcinoma	Lung 16	2.58	0.37
5	Lng SQ45	Squamous cell carcinoma	Lung 17	4.61	1.53
	Lng 315L	Squamous cell carcinoma	Lung 18	1.15	1.16
	Lng SQ14	Squamous cell carcinoma	Lung 19	1.83	0.78
	Lng SQ9X	Squamous cell carcinoma	Lung 20	2.70	0.14
	Lng SQ56	Squamous cell carcinoma	Lung 21	2.50	1.53
10	Lng SQ80	Squamous cell carcinoma	Lung 22	2.69	0.77
	Lng SQ32	Squamous cell carcinoma	Lung 23	7.70	1.51
	Lng SQ16	Squamous cell carcinoma	Lung 24	0.70	0.04
	Lng SQ79	Squamous cell carcinoma	Lung 25	3.61	0.92
	Lng 90X	Squamous cell carcinoma	Lung 26	1.24	0.23
15	Lng 47XQ	Squamous cell carcinoma	Lung 27	1.90	0.13
	Lng BR94	Squamous cell carcinoma	Lung 28	2.87	0.00
	Lng C20X	Squamous cell carcinoma	Lung 29	0.05	0.04
	Lng SQ44	Squamous cell carcinoma	Lung 30	0.21	2.13
	Lng SQ43	Squamous cell carcinoma	Lung 31	2.86	0.04
20	Lng LC71	Large cell carcinoma	Lung 32	1.94	1.82
	Lng LC109	Large cell carcinoma	Lung 33	4.04	4.30
	Lng LC80	Large cell carcinoma	Lung 34	6.13	0.51
	Lng 77L	Large cell carcinoma	Lung 35	0.03	1.08
	Lng 75XC	Metastatic from bone cancer	Lung 36	0.15	0.19
25	Lng MT71	Metastatic from renal cell cancer	Lung 37	5.96	0.74
	Lng MT67	Metastatic from melanoma	Lung 38	12.30	1.18
	Bld46XK		Bladder 1	0.03	0.02
	BldTR14		Bladder 2	2.89	1.51
	Blo B5		Blood 1		21.19
30	Blo B6		Blood 2		41.21
	Cvx KS52		Cervix 1	5.78	1.44
	Cvx KS83		Cervix 2	17.75	4.29
	ClnAS43		Colon 1	3.42	0.10
	ClnAS45		Colon 2	0.17	0.13
35	ClnAS46		Colon 3	2.29	1.92
	ClnAS67		Colon 4	0.20	0.33
	ClnAS89		Colon 5	0.08	0.12
	End 10479		Endometrium	25.63	4.63
			1		

00000000000000000000000000000000

	End 28XA	Endometrium	6.25	2.46
	End 68X	Endometrium	6.43	11.24
	Kid10XD	Kidney 1	3.73	1.07
5	Kid 109XD	Kidney 2	2.90	4.82
	Liv15XA	Liver 1	0.19	0.08
	liv 174 L	Liver 2	0.99	0.76
	Mam 173 M	Mammary 1	0.76	0.47
10	Mam 220	Mammary 2	0.11	0.23
	Mam 355	Mammary 3	1.08	0.19
	Mam 976M	Mammary 4	0.02	0.16
	ovr 180B	Ovary 1	16.11	
	Ovr 18GA	Ovary 2	15.14	
	Ovr A084	Ovary 3	8.06	5.58
15	Pan 77X	Pancreas 1	3.94	2.51
	Pan 92X	Pancreas 2	4.98	1.70
	Pro 101XB	Prostate 1	3.69	2.15
	Pro 109XB	Prostate 2	0.16	0.23
	Pro 125XB	Prostate 3	0.11	0.12
20	Pro 13XB	Prostate 4	0.04	0.31
	Skn 39A	Skin 1	1.19	0.08
	Skn 816S	Skin 2	0.95	0.01
	SmInt	Small	0.75	0.10
	21XA	Intestine 1		
	SmInt H89	Small	0.66	0.46
		Intestine 2		
25	Sto 115S	Stomach 1	1.91	1.20
	Sto 264S	Stomach 2	0.74	0.99
	Sto288S	Stomach 3	2.78	0.06
	Tst647T	Testis 1	1.87	2.68
30	Tst 663T	Testis 2	7.89	0.66
	Thr 270T	Thyroid 1	2.01	2.13
	Thr 939T	Thyroid 2	0.50	0.55
	Utr135XO	Uterus 1	3.52	6.06
	Utr 141XO	Uterus 2	2.59	2.57

0= Negative

35 In the analysis of matching samples, higher expression of Lng143 is detected in lung samples showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal pooled samples (Table 1).

40 Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher levels of mRNA expression in the cancer sample compared to 45 the normal adjacent). Table 2 shows differential expression of Lng143 in 38 lung cancer tissues compared with their respective normal adjacent tissue.

Altogether, the high level of tissue specificity, plus the mRNA differential expression in the lung matching samples tested are believed to make Lng143 a good marker for diagnosing, monitoring, staging, imaging and treating 5 lung cancer.

**DNA sequence for Lng143**

Sequence available from Incyte database.

**Primers Used for QPCR Expression Analysis****Forward primer**

10 CCGACCTTGAGATTATTCCTGT (SEQ ID NO:47)

**Reverse primer**

GCACCACTTAAACCAAATCCA (SEQ ID NO:48)

**Probe**

TGCTGCCAACACCACTTCTCCATCT (SEQ ID NO:49)

**15 Example 11****Sequence 11****Lng144**

Gene ID 52017

ddx lung code SQlng007

20 **Table 1.** The absolute numbers are relative levels of expression of Lng144 in 24 normal different tissues. All the values are compared to normal uterus (calibrator). These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from 25 different individuals.

Tissue	NORMAL
Adrenal Gland	0.04
Bladder	1.29
Brain	0.44
Cervix	0.85
Colon	0.00
Endometrium	0.43

	Esophagus	0.05
	Heart	0.06
	Kidney	0.18
	Liver	0.30
5	Lung	1.35
	Mammary Gland	1.04
	Muscle	0.34
	Ovary	0.29
	Pancreas	0.77
10	Prostate	0.93
	Rectum	0.26
	Small	0.11
	Spleen	3.92
	Stomach	0.30
15	Testis	1.1
	Thymus	0.93
	Trachea	0.69
	Uterus	1.00

0=negative

20 The relative levels of expression in Table 1 show that Lng144 mRNA expression is high in lung compared with other normal tissues analyzed.

25 The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

Table 2. The absolute numbers are relative levels of expression of Lng144 in 30 pairs of matching samples. All 30 the values are compared to normal uterus (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING NORMAL ADJACENT
35	Lng 60L	Adenocarcinoma	Lung 1	0.65	0.30
	Lng 143L	Adenocarcinoma	Lung 2	0.29	0.17
	Lng AC66	Adenocarcinoma	Lung 3	0	0.34
	Lng AC69	Adenocarcinoma	Lung 4	1.41	1.89
40	Lng AC11	Adenocarcinoma	Lung 5	2.82	3.11
	Lng AC32	Adenocarcinoma	Lung 6	1.09	1.27
	Lng AC94	Adenocarcinoma	Lung 7	2.20	0.84

5	Lng 223L	Adenocarcinoma	Lung 8	0.30	0.27
	Lng BR26	Bronchio-alveolar carcinoma	Lung 9	1.25	0.36
	Lng SQ45	Squamous cell carcinoma	Lung 10	2.96	1.26
	Lng SQ9X	Squamous cell carcinoma	Lung 11	1.49	0.30
5	Lng SQ80	Squamous cell carcinoma	Lung 12	1.88	1.51
	Lng SQ16	Squamous cell carcinoma	Lung 13	0.48	0.47
	Lng SQ79	Squamous cell carcinoma	Lung 14	2.77	0.00
	Lng 90X	Squamous cell carcinoma	Lung 15	0.09	0.29
10	Lng SQ43	Squamous cell carcinoma	Lung 16	0.81	0.26
	Lng LC71	Large cell Carcinoma I	Lung 17	0.98	1.65
	Lng LC109	Large cell carcinoma IIIA	Lung 18	0.38	1.63
	Lng MT71	Metastatic from melanoma	Lung 19	0.13	0.20
	Lng MT67	Metastatic from renal cancer	Lung 20	0.63	0.77
15	Bld46XK		Bladder 1	0.24	0.08
	BldTR14		Bladder 2	0.18	0.95
	ClnAS45		Colon2	0.12	0.05
	ClnAS46		Colon3	0.21	0.98
	ClnAS67		Colon4	0.09	0.18
20	ClnAS89		Colon 5	0.38	3.31
	ClnAS43		Colon5	0.18	0.47
	Liv15XA		Liver 1	0.47	0.05
	Tst647T		Testis 1	2.10	0.26
	Utr135XO		Uterus 1	0.81	0.80

0= Negative

25        In the analysis of matching samples, higher expression of lng144 is detected in lung samples showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal pooled samples (Table 1).

30        Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher levels of mRNA expression in the cancer sample compared to 35 the normal adjacent). Table 2 shows differential expression of Lng144 in 20 lung cancer tissues compared with their respective normal adjacent tissue.

Altogether, the high level of tissue specificity, plus the mRNA differential expression in the lung matching samples tested are believed to make Lng144 a good marker for diagnosing, monitoring, staging, imaging and treating 5 lung cancer.

**DNA sequence for Lng144**

Sequence available from Incyte database.

**Primers Used for QPCR Expression Analysis****Forward primer**

10 TGCTGCCACAAACCGAGA (SEQ ID NO:50)

**Reverse primer**

TTGGGAGGGTTGGTTGGTT (SEQ ID NO:51)

**Probe**

TTTGAGGGCACTAGGAAACGATCTGT (SEQ ID NO:52)

**15 Example 12****Sequence 12****Lng138**

Gene ID 460254

ddx lung code SQLng110

20 **Table 1.** The absolute numbers are relative levels of expression of Lng138 in 24 normal different tissues. All the values are compared to normal spleen (calibrator). These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from 25 different individuals.

Tissue	NORMAL
Adrenal Gland	0.00
Bladder	0.00
Brain	0.04
Cervix	0.09
Colon	0.00
Endometrium	0.31

	Esophagus	0.00
	Heart	0.00
	Kidney	0.03
	Liver	0.00
5	Lung	1.35
	Mammary Gland	0.02
	Muscle	0.03
	Ovary	1.15
	Pancreas	0.03
10	Prostate	0.27
	Rectum	0.02
	Small Intestine	0.02
	Spleen	1.00
	Stomach	0.05
15	Testis	0.08
	Thymus	1.00
	Trachea	0.12
	Uterus	0.24

0=negative

20 The relative levels of expression in Table 1 show that Lng138 mRNA expression is high in lung compared with other normal tissues analyzed.

25 The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

30 **Table 2.** The absolute numbers are relative levels of expression of Lng138 in 50 pairs of matching samples. All the values are compared to normal spleen (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING NORMAL ADJACENT
35	Lng 60L	Adenocarcinoma	Lung 1	1.66	0.82
	Lng 143L	Adenocarcinoma	Lung 2	0.06	0.03
	Lng AC82	Adenocarcinoma	Lung 3	0.02	0.03
	Lng AC66	Adenocarcinoma	Lung 4	0.1	0.79
	Lng AC69	Adenocarcinoma	Lung 5	0.56	0.65
	Lng 60XL	Adenocarcinoma	Lung 6	0.02	0.15
	Lng AC94	Adenocarcinoma	Lung 7	0.67	0.65
	Lng AC11	Adenocarcinoma	Lung 8	0.41	3.32
	Lng AC32	Adenocarcinoma	Lung 9	0.47	1.57

	Lng 47XQ	Adenocarcinoma	Lung 10	0.03	0.26
	Lng 223L	Adenocarcinoma	Lung 11	0.3	0.16
	Lng BR26	Bronchogenic carcinoma	Lung 12	0.12	0.26
	Lng BA641	Bronchio-alveolar carcinoma	Lung 13	0.17	0.1
5	LngSQ45	Squamous cell carcinoma	Lung 14	5.3	0.22
	Lng SQ14	Squamous cell carcinoma	Lung 15	0.02	0.4
	Lng SQ9X	Squamous cell carcinoma	Lung 16	0.38	0.06
	Lng SQ56	Squamous cell carcinoma	Lung 17	0.05	0.73
	Lng SQ16	Squamous cell carcinoma	Lung 18	0.04	0.4
	Lng SQ32	Squamous cell carcinoma	Lung 19	0.11	0.68
10	Lng SQ80	Squamous cell carcinoma	Lung 20	0.17	0.38
	Lng SQ79	Squamous cell carcinoma	Lung 21	0.29	0.79
	Lng SQ43	Squamous cell carcinoma	Lung 22	0.19	0.01
	Lng BR94	Squamous cell carcinoma	Lung 23	0.1	0
	Lng 90X	Squamous cell carcinoma	Lung 24	0.01	0.02
	Lng LC71	Large cell carcinoma	Lung 25	0.2	0.47
15	LC109	Large cell carcinoma	Lung 26	0.09	0.54
	Lng MT67	Metastatic from renal carcinoma	Lung 27	0.4	0.11
	Bld46XK		Bladder 1	0	0
	BldTR14		Bladder 2	0.1	0.04
	ClnAS43		Colon 1	0.11	0.06
	ClnAS45		Colon 2	0.03	0.03
20	ClnAS46		Colon 3	0.08	0.04
	ClnAS67		Colon 4	0.05	0.11
	ClnAS89		Colon 5	0.02	0.08
	End 28XA		Endometrium	0.08	0.12
	Kid 109XD		1		
			Kidney 1	0.12	0.05
25	Kid 10XD		Kidney 2	0.09	0.02
	Liv15XA		Liver 1	0.01	0
	Mam 173M		Mammary 1	0.03	0.02
	Mam 220		Mammary 2	0.01	0.26
	Mam 355		Mammary 3	0.1	0.03
	Ovr A084		Ovary 1	0.11	0.08
30	Pro 101XB		Prostate 1	0.07	0.09
	Pro 109XB		Prostate 2	0.01	0.01
	Pro 125XB		Prostate 3	0	0.01
	Sto 115S		Stomach 1	0.07	0.07
	Sto 246S		Stomach 2	0.16	0.1
	Sto 288S		Stomach 3	0.07	0.01
35	Tst 647T		Testis 2	0.13	0.02
	Utr135XO		Uterus 1	0.41	0.18

0= Negative

In the analysis of matching samples, higher expression of *lnc138* is detected in lung samples showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal 5 pooled samples (Table 1).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows differential expression of Lng138 in 27 lung cancer tissues compared with their respective normal adjacent tissue.

15 Altogether, the high level of tissue specificity, plus the mRNA differential expression in the lung matching samples tested are believed to make Lng138 a good marker for diagnosing, monitoring, staging, imaging and treating lung cancer.

### DNA sequence for Lng138

20 Sequence available from Incyte database.

#### Primers Used for QPCR Expression Analysis

### Forward primer

CCTGATACCTTAACCAATGCTCT (SEQ ID NO:53)

### Reverse primer

25 TTGGGTAGTATCAAATGGGTAAGG (SEQ ID NO:54)

## Probe

CCTGTCCTTCTCCTTGCGTTATGCTATCC (SEQ ID NO:55)

### Example 13

### Sequence 13

30 Lng137

Gene ID 179090

ddx lung code SQLng012

**Table 1.** The absolute numbers are relative levels of expression of Lng137 in 24 normal different tissues. All the values are compared to normal spleen (calibrator).

5 These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

	Tissue	NORMAL
10	Adrenal Gland	0.042
	Bladder	0.063
	Brain	0.285
	Cervix	0.196
	Colon	0.080
15	Endometrium	0.956
	Esophagus	0.025
	Heart	0.010
	Kidney	0.046
	Liver	0.035
20	Lung	0.204
	Mammary Gland	0.142
	Muscle	0.092
	Ovary	0.760
	Pancreas	0.084
25	Prostate	0.355
	Rectum	0.357
	Small Intestine	0.074
	Spleen	1.000
	Stomach	0.103
	Testis	2.612
30	Thymus	10.853
	Trachea	0.076
	Uterus	0.235

0=negative

The relative levels of expression in Table 1 show that  
35 Lng137 mRNA expression is relatively high in lung compared with most other normal tissues analyzed.

The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the  
40 absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

**Table 2.** The absolute numbers are relative levels of expression of Lng137 in 70 pairs of matching samples. All the values are compared to normal spleen (calibrator). A matching pair is formed by mRNA from the cancer sample for 5 a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING
10	Lng60L	Adenocarcinoma	Lung 1	0.92	0.67
	Lng 143L	Adenocarcinoma	Lung 2	0.53	0.04
	Lng AC82	Adenocarcinoma	Lung 3	2.7	0
	LngAC66	Adenocarcinoma	Lung 4	1.62	0.34
	LngAC69	Adenocarcinoma	Lung 5	3.18	0.79
15	Lng AC88	Adenocarcinoma	Lung 6	1.87	0.32
	Lng 60XL	Adenocarcinoma	Lung 7	3.42	0.24
	LngAC94	Adenocarcinoma	Lung 8	0	0.21
	LngAC11	Adenocarcinoma	Lung 9	23.43	2.76
	LngAC32	Adenocarcinoma	Lung 10	5.17	0.63
	Lng 47XQ	Adenocarcinoma	Lung 11	2.03	0
	Lng AC90	Adenocarcinoma	Lung 12	4.69	0
20	Lng AC39	Adenocarcinoma	Lung 13	2.48	0.22
	Lng223L	Adenocarcinoma	Lung 14	1.89	0
	Lng 528L	Adenocarcinoma	Lung 15	1.47	0
	Lng BR26	Bronchio-alveolar	Lung 16	13.18	0.47
	Lng BA641	Squamous cell	Lung 17	0.97	0.18
25	Lng 315L	Squamous cell	Lung 18	0.63	0.62
	LngSQ45	Squamous cell	Lung 19	3.2	18.7
	Lng SQ14	Squamous cell	Lung 20	2.28	0.29
	Lng SQ9X	Squamous cell	Lung 21	1	0.06
	Lng SQ56	Squamous cell	Lung 22	10.27	0.67
30	LngSQ16	Squamous cell	Lung 23	2.06	0
	Lng SQ32	Squamous cell	Lung 24	1.34	0.58
	Lng SQ80	Squamous cell	Lung 25	5.21	0.29
	LngSQ79	Squamous cell	Lung 26	10.59	0.93
	Lng C20X	Squamous cell	Lung 27	0	0
35	Lng SQ43	Squamous cell	Lung 28	16.56	0.04
	Lng BR94	Squamous cell	Lung 29	9.92	0
	Lng SQ44	Squamous cell	Lung 30	0.13	6.23
	Lng 90X	Squamous cell	Lung 31	0.31	0
	Lng 77L	Large cell carcinoma	Lung 32	0.07	1.75
40	Lng LC71	Large cell carcinoma	Lung 33	1.93	0.55
	Lng LC109	Large cell carcinoma	Lung 34	13.98	0.79
	Lng LC80	Large cell carcinoma	Lung 35	7.31	0
	Lng MT67	Metastatic from bone	Lung 36	2.03	0.17
	Lng MT71	Metastatic from renal	Lung 37	1.55	0
45	Lng 75XC	Metastatic from	Lung 38	0.04	0.17
	Bld46XK		Bladder 1	0	0
	BldTR14		Bladder 2	4.47	2.79

	Cvx KS52	Cervix 1	4	0.53
	Cvx KS83	Cervix 2	6.43	0.92
	ClnAS43	Colon 1	12.55	0.77
	ClnAS45	Colon 2	1.51	0.35
5	ClnAS46	Colon 3	18	1.6
	ClnAS67	Colon 4	1.6	5.48
	ClnAS89	Colon 5	0.78	0.04
	End 28XA	Endometrium	3.66	0.64
	End 8911	Endometrium	1.68	2.36
10	Kid 109XD	Kidney 1	0.52	0.44
	Kid 10XD	Kidney 2	0.51	0.07
	Liv15XA	Liver 1	0	0.02
	Liv174L	Liver 2	0.16	0.09
	Mam 162X	Mammary 1	0.03	0.1
15	Mam 355	Mammary 2	0.13	0.08
	Ovr A084	Ovary 1	1.42	1.3
	Pan 71XL	Pancreas 1	1.01	0.38
	Pan 92X	Pancreas 2	0.64	1.16
	Pro 109XB	Prostate 1	0.08	0.15
20	Pro 125XB	Prostate 2	0.03	0.03
	SmInt	Small	0.33	0.03
	SmInt H89	Small	2.94	0.12
	Sto 758S	Stomach 1	2.89	0.13
	Sto 288S	Stomach 2	1.03	0.05
25	Tst 47T	Testis 1	3.93	0.56
	Tst 39X	Testis 2	1.9	1.39
	Thr 143N	Thyroid 1	0.16	0.36
	Thr 270T	Thyroid 2	0.35	0
	Utr 35XO	Uterus 1	2.9	2.62
30	Utr 141XO	Uterus 2	0.87	0.38

0 = Negative

In the analysis of matching samples, higher expression of lng137 are detected in lung samples showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal pooled samples (Table 1).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows overexpression of Lng137 in 31 lung cancer tissues compared with their respective normal adjacent tissue in 38 cancer matching pairs (lung samples #2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13,

14, 15, 16, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 31, 33, 34, 35, 36, and 37). There is overexpression in the cancer tissue for 82% of the lung matching samples tested (total of 38 lung matching samples).

5 Altogether, the high level of tissue specificity, plus the mRNA overexpression in 82% of the lung matching samples tested are believed to make Lng137 a good marker for diagnosing, monitoring, staging, imaging and treating lung cancer.

10 **Northern Analysis**

One transcript ~ 3.4 kb

**DNA sequence for Lng137**

Sequence available from Incyte database.

**Primers Used for QPCR Expression Analysis**

15 **Forward primer**  
CTCGGATATGATTAAAGAGTTTCG (SEQ ID NO:56)  
**Reverse primer**  
TCCACTGTGCTGTTGTT (SEQ ID NO:57)  
**Probe**  
20 ATTGGCGTGCTCTTGTAACCTCTGAGA (SEQ ID NO:58)

**Example 14**

**Sequence 14**

**Lng142**

Gene ID 6348

25 ddxlung code SQLng004

**Table 1.** The absolute numbers are relative levels of expression of Lng142 in 24 normal different tissues. All the values are compared to normal lung (calibrator). These RNA samples are commercially available pools, originated by 30 pooling samples of a particular tissue from different individuals.

Tissue	NORMAL
Adrenal Gland	0.02
Bladder	0.00
Brain	0.00
5 Cervix	0.00
Colon	0.00
Endometrium	0.00
Esophagus	0.00
Heart	0.00
10 Kidney	0.00
Liver	0.00
Lung	1.00
Mammary Gland	0.01
15 Muscle	0.00
Ovary	0.03
Pancreas	0.01
Prostate	0.02
Rectum	0.00
20 Small Intestine	0.01
Spleen	0.00
Stomach	0.00
Testis	0.03
Thymus	0.00
Trachea	0.01
25 Uterus	0.03

0=negative

The relative levels of expression in Table 1 show that Lng142 mRNA expression is high in lung compared with most other normal tissues analyzed.

30 The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

35 **Table 2.** The absolute numbers are relative levels of expression of Lng142 in 20 pairs of matching samples. All the values are compared to normal lung (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent 40 sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING NORMAL ADJACENT
5	Lng 60L	Adenocarcinoma	Lung 1	0.38	0.22
	Lng AC66	Adenocarcinoma	Lung 2	0.00	0.05
	Lng AC69	Adenocarcinoma	Lung 3	0.00	0.00
	Lng AC11	Adenocarcinoma	Lung 4	0.00	1.37
	Lng AC32	Adenocarcinoma	Lung 5	0.05	0.16
	Lng AC94	Adenocarcinoma	Lung 6	0.08	0.00
	Lng 223L	Adenocarcinoma	Lung 7	0.03	0.00
	Lng SQ45	Squamous cell carcinoma	Lung 8	3.82	0.02
10	Lng SQ16	Squamous cell carcinoma	Lung 9	0.00	0.00
	Lng SQ79	Squamous cell carcinoma	Lung 10	0.00	0.11
15	Bld 46XK		Bladder 1	0.00	0.00
	Bld TR14		Bladder 2	0.00	0.00
	Cln AS45		Colon 1	0.00	0.00
	Cln AS46		Colon 2	0.00	0.00
	Cln AS67		Colon 3	0.00	0.01
	Cln AS89		Colon 4	0.01	0.02
20	Cln AS43		Colon 5	0.00	0.00
	Liv 15XA		Liver 1	0.00	0.00
	Tst 647T		Testis 1	0.00	0.05
	Utr 135XO		Uterus 1	0.17	0.00

0 = Negative

In the analysis of matching samples, higher expression of lng142 is detected in lung samples showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal pooled samples (Table 1).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows differential expression of Lng142 in 10 lung cancer tissues compared with their respective normal adjacent tissue.

Altogether, the high level of tissue specificity, plus the mRNA differential expression in the lung matching samples tested are believed to make Lng142 a good marker for diagnosing, monitoring, staging, imaging and treating 5 lung cancer.

**DNA sequence for Lng142**

Sequence available from Incyte database.

**Primers Used for QPCR Expression Analysis****Forward primer**

10 TGGCTAAAATAGGTCTTGTAGGGA (SEQ ID NO:59)

**Reverse primer**

CAAGGAGGGGGCATTGTA (SEQ ID NO:60)

**Probe**

TCCTTCCTTGGCAATCTCCTCTCCTG (SEQ ID NO:61)

**15 Example 15****Sequence 15****Lng140**

Gene ID 94694

ddx lung code SQLng005

20 **Table 1.** The absolute numbers are relative levels of expression of Lng140 in 24 normal different tissues. All the values are compared to normal mammary gland (calibrator). These RNA samples are commercially available pools, originated by pooling samples of a particular tissue 25 from different individuals.

Tissue	NORMAL
Adrenal Gland	0
Bladder	0.00
Brain	0.00
Cervix	0.00
Colon	0.00
Endometrium	0.14

	Esophagus	0.00
	Heart	0.00
	Kidney	0.00
	Liver	0.00
5	Lung	183.55
	Mammary Gland	1.00
	Muscle	0.00
	Ovary	0.00
	Pancreas	0.00
10	Prostate	0.10
	Rectum	0.06
	Small	0.03
	Spleen	0.00
	Stomach	0.02
15	Testis	0.01
	Thymus	0.00
	Trachea	3.72
	Uterus	0.00

0=negative

20 The relative levels of expression in Table 1 show that Lng140 mRNA expression is much higher in lung compared with most other normal tissues analyzed.

25 The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

Table 2. The absolute numbers are relative levels of expression of Lng140 in 78 pairs of matching samples, 2 blood samples, 1 normal ovary and 1 cancer ovary sample. All the values are compared to normal mammary gland (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

40	Sample ID	Cancer Type	Tissue	CANCER	MATCHING	
					NORMAL	NORMAL ADJACENT
	Lng 60L	Adenocarcinoma	Lung 1	21.56	3.54	
	Lng 143L	Adenocarcinoma	Lung 2	0.00	7.31	
	Lng 60XL	Adenocarcinoma	Lung 3	88.03	19.84	
	Lng AC82	Adenocarcinoma	Lung 4	4.61	122.36	
	Lng AC88	Adenocarcinoma	Lung 5	4.61	70.77	

	Lng AC66	Adenocarcinoma	Lung 6	38.19	7.46	
	Lng AC69	Adenocarcinoma	Lung 7	36.00	25.99	
	Lng AC11	Adenocarcinoma	Lung 8	1287.18	280.14	
5	Lng AC32	Adenocarcinoma	Lung 9	208.66	0.00	
	Lng AC94	Adenocarcinoma	Lung 10	39.81	153.28	
	Lng AC90	Adenocarcinoma	Lung 11	78.25	420.22	
	Lng AC39	Adenocarcinoma	Lung 12	600.49	5.17	
	Lng 223L	Adenocarcinoma	Lung 13	5.60	3.56	
10	Lng 528L	Adenocarcinoma	Lung 14	6.17	28.05	
	Lng BR26	Bronchogenic carcinoma	Lung 15	4.68	23.18	
	Lng BA641	Bronchio-alveolar carcinoma	Lung 16	263.20	4.11	
	Lng 315L	Squamous cell carcinoma	Lung 17	0.00	3.77	
15	Lng SQ14	Squamous cell carcinoma	Lung 18	0.74	5.60	
	Lng SQ56	Squamous cell carcinoma	Lung 19	36.25	186.75	
	Lng SQ9X	Squamous cell carcinoma	Lung 20	77.98	1.99	
	Lng SQ80	Squamous cell carcinoma	Lung 21	20.32	35.02	
	Lng SQ45	Squamous cell carcinoma	Lung 22	153.28	80.73	
	Lng SQ16	Squamous cell carcinoma	Lung 23	9.45	13.04	
20	Lng SQ32	Squamous cell carcinoma	Lung 24	3213.66	99.04	
	Lng SQ79	Squamous cell carcinoma	Lung 25	594.28	48.17	
	Lng 47XQ	Squamous cell carcinoma	Lung 26	47.84	0.00	
	Lng BR94	Squamous cell carcinoma	Lung 27	4.66	0.00	
	Lng 90X	Squamous cell carcinoma	Lung 28	0.00	6.41	
25	Lng C20X	Squamous cell carcinoma	Lung 29	2.35	0.00	
	Lng SQ44	Squamous cell carcinoma	Lung 30	6.59	1.55	
	Lng SQ43	Squamous cell carcinoma	Lung 31	25.19	0.00	
	Lng LC71	Large cell carcinoma	Lung 32	1408.55	97.01	
30	Lng LC109	Large cell carcinoma	Lung 33	85.92	922.88	
	Lng LC80	Large cell carcinoma	Lung 34	99.39	11.16	
	Lng 77L	Large cell carcinoma	Lung 35	8.69	11.35	
	Lng 75XC	Metastatic from bone cancer	Lung 36	0.00	0.00	
	Lng MT67	Metastatic from renal cell cancer	Lung 37	0.00	2.28	
35	Lng MT71	Metastatic from melanoma	Lung 38	1.56	0.00	
	Bld46XK		Bladder 1	0.00	0.00	
	BldTR14		Bladder 2	0.00	168.90	0.00
	Blo B5		Blood 1			0.00
	Blo B6		Blood 2			0.00
40	Cvx KS52		Cervix 1	85.33	0.00	

5	Cvx KS83	Cervix 2	23.51	0.00
	ClnAS43	Colon1	259.57	0.00
	ClnAS45	Colon2	0.00	0.00
	ClnAS46	Colon3	14.52	0.00
	ClnAS67	Colon4	5.90	41.64
	ClnAS89	Colon5	5.13	6.54
	End	Endometri	0.00	0.00
	10479	um 1		
	End 28XA	Endometri	38.45	1.29
	um 2			
10	End 68X	Endometri	0.00	2.30
	Kid10XD	um 3		
	Kid	Kidney 1	0.00	0.00
	109XD	Kidney 2	0.00	0.00
	Liv15XA	Liver 1	0.00	0.00
15	liv 174	Liver 2	0.00	0.00
	L			
	Mam 173	Mammary 1	0.87	0.00
	M			
	Mam 220	Mammary 2	0.00	0.00
20	Mam 355	Mammary 3	0.00	0.00
	Mam 976M	Mammary 4	0.00	0.00
	ovr 180B	ovary 1	0.00	0.00
	Ovr 18GA	Ovary 2		0.00
	Ovr A084	Ovary 3	36.50	
25	Pan 77X	Pancreas	0.00	0.00
	Pan 92X	1		
		Pancreas	46.53	0.00
	Pro	2		
	101XB	Prostate	0.29	1.43
30	Pro	1		
	109XB	Prostate	0.00	0.00
	Pro	2		
	125XB	Prostate	1.30	1.97
	Pro 13XB	3		
35		Prostate	0.00	0.00
	4			
	Skn 39A	Skin 1	0.00	0.00
	Skn 816S	Skin 2	0.00	0.00
	SmInt	Small	1.45	1.70
40	21XA	Intestine		
	1	1		
	SmInt	Small	79.07	1.17
	H89	Intestine		
	2	2		
45	Sto 115S	Stomach 1	109.14	10.16
	Sto 264S	Stomach 2	2.53	0.00
	Sto288S	Stomach 3	0.00	0.00
	Tst647T	Testis 1	0.00	0.00
	Tst 663T	Testis 2	0.00	0.00
50	Thr 270T	Thyroid 1	0.00	0.00
	Thr 939T	Thyroid 2	0.00	0.00
	Utr135XO	Uterus 1	3.89	0.00
	Utr	Uterus 2	0.29	1.43
	141XO			

50 0= Negative

In the analysis of matching samples, the higher expression level of Lng140 is detected in lung samples showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity 5 results obtained with normal pooled samples (Table 1).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher 10 levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows overexpression of Lng140 in 20 lung cancer tissues compared with their respective normal adjacent tissue in 38 cancer matching pairs (lung samples #1, 3, 6, 8, 9, 12, 13, 16, 20, 22, 24- 15 27, 29-32, 34, and 38). There is overexpression in the cancer tissue for 53% of the lung matching samples tested (total of 38 lung matching samples).

Altogether, the high level of tissue specificity, plus the mRNA overexpression in 53% of the lung matching samples 20 tested are believed to make Lng140a good marker for diagnosing, monitoring, staging, imaging and treating lung cancer.

#### Primers Used for QPCR Expression Analysis

##### **Forward primer**

25 CCTCTGAAGAACGATCACAAACA (SEQ ID NO:62)

##### **Reverse primer**

ATTCCAGCCTGAGTCACACAGA (SEQ ID NO:63)

##### **Probe**

ACCAAGGAGAAACAAACCAAGCAGCA (SEQ ID NO:64)

#### 30 Example 16

##### **Sequence 16**

**Lng151**

Gene ID 145812

ddxlung code SQlng008

**Table 1.** The absolute numbers are relative levels of expression of Lng151 in 24 normal different tissues. All the values are compared to normal thymus (calibrator). These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

	Tissue	NORMAL
10	Adrenal Gland	0.01
	Bladder	0.01
	Brain	0.03
15	Cervix	0.07
	Colon	0.01
	Endometrium	0.16
	Esophagus	0.02
	Heart	0.00
20	Kidney	0.01
	Liver	0.00
	Lung	0.17
	Mammary Gland	0.06
	Muscle	0.04
25	Ovary	0.44
	Pancreas	0.05
	Prostate	0.04
	Rectum	0.03
	Small Intestine	0.01
	Spleen	0.13
30	Stomach	0.02
	Testis	0.03
	Thymus	1.00
	Trachea	0.09
35	Uterus	0.09

0=negative

The relative levels of expression in Table 1 show that Lng151 mRNA expression is high in lung compared with most other normal tissues analyzed.

The absolute numbers in Table 1 were obtained 40 analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the

absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

**Table 2.** The absolute numbers are relative levels of expression of Lng151 in 20 pairs of matching samples. All 5 the values are compared to normal thymus (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING NORMAL ADJACENT
10	Lng 60L	Adenocarcinoma	Lung 1	0.35	0.18
	Lng AC66	Adenocarcinoma	Lung 2	0.37	0.34
	Lng AC69	Adenocarcinoma	Lung 3	1.99	0.32
	Lng AC11	Adenocarcinoma	Lung 4	1.13	1.11
	Lng AC32	Adenocarcinoma	Lung 5	0.75	0.23
	Lng AC94	Adenocarcinoma	Lung 6	0.2	0.1
15	Lng 223L	Adenocarcinoma	Lung 7	0.06	0
	Lng SQ45	Squamous cell carcinoma	Lung 8	2.45	0.94
	Lng SQ16	Squamous cell carcinoma	Lung 9	0.18	0.05
	Lng SQ79	Squamous cell carcinoma	Lung 10	1.23	0.62
20	Bld 46XK		Bladder 1	0.06	0
	Bld TR14		Bladder 2	0.27	0.37
	Cln AS43		Colon 5	0.27	0.04
	Cln AS45		Colon 1	0.02	0.04
	Cln AS46		Colon 2	0.04	0.15
25	Cln AS67		Colon 3	0.03	0.28
	Cln AS89		Colon 4	0.05	0.32
	Liv 15XA		Liver 1	0.22	0.08
	Tst 647T		Testis 1	0.26	0.21
	Utr 135XO		Uterus 1	1.14	1.06

30 0= Negative

In the analysis of matching samples, higher expression of lng151 is detected in lung samples showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal 35 pooled samples (Table 1).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher 40 levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows differential

expression of Lng151 in 10 lung cancer tissues compared with their respective normal adjacent tissue.

Altogether, the high level of tissue specificity, plus the mRNA differential expression in the lung matching 5 samples tested are believed to make Lng151 a good marker for diagnosing, monitoring, staging, imaging and treating lung cancer.

**DNA sequence for Lng151**

Sequence available from Incyte database.

10 **Primers Used for QPCR Expression Analysis**

**Forward primer**

TGAGGAGAAAGAAGGGAATCAC (SEQ ID NO:65)

**Reverse primer**

TCCTAAGGTAGCACTATTGGAGAC (SEQ ID NO:66)

15 **Probe**

AGCAATGAAGAATGAACCTGGAGTAAAGAGTCA (SEQ ID NO:67)

**Example 17**

**Sequence 17**

**Lng150**

20 Gene ID 10713

ddx lung code SQLng002

**Table 1.** The absolute numbers are relative levels of expression of Lng150 in 24 normal different tissues. All the values are compared to normal testis (calibrator).

25 These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

Tissue	NORMAL
Adrenal Gland	0.00
Bladder	0.04
Brain	0.01

	Cervix	0.00
	Colon	0.00
	Endometrium	0.04
	Esophagus	0.00
5	Heart	0.00
	Kidney	0.01
	Liver	0.00
	Lung	0.01
10	Mammary Gland	0.00
	Muscle	0.00
	Ovary	0.00
	Pancreas	0.00
	Prostate	0.03
	Rectum	0.00
15	Small Intestine	0.00
	Spleen	0.00
	Stomach	0.00
	Testis	1.00
	Thymus	0.00
20	Trachea	0.01
	Uterus	0.07

0=negative

The relative levels of expression in Table 1 show that Lng150 mRNA expression is detected in lung and is not 25 detectable in adrenal gland, cervix, colon, esophagus, heart, liver, mammary gland, muscle, ovary, pancreas, rectum, small intestine, spleen, stomach, and thymus.

The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from 30 different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

**Table 2.** The absolute numbers are relative levels of expression of Lng150 in 40 pairs of matching samples. All 35 the values are compared to normal testis (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

40	Sample ID	Cancer Type	Tissue	CANCER	MATCHING NORMAL ADJACENT
	Lng 60L	Adenocarcinoma	Lung 1	0.00	0.01
	Lng AC88	Adenocarcinoma	Lung 2	0.02	0.00

	Lng AC66	Adenocarcinoma	Lung 3	0.00	0.24
	Lng AC69	Adenocarcinoma	Lung 4	0.02	0.00
	LngAC11	Adenocarcinoma	Lung 5	0.00	0.20
	Lng AC32	Adenocarcinoma	Lung 6	0.04	0.00
5	Lng AC39	Adenocarcinoma	Lung 7	0.00	0.00
	Lng AC94	Adenocarcinoma	Lung 8	0.00	0.00
	Lng AC90	Adenocarcinoma	Lung 9	0.00	0.00
	Lng 223L	Adenocarcinoma	Lung 10	0.00	0.00
	Lng BR26	Bronchio-alveolar carcinoma	Lung 11	0.02	0.02
10	Lng SQ45	Bronchogenic carcinoma	Lung 12	1.08	0.04
	Lng SQ9X	Squamous cell carcinoma	Lung 13	0.00	0.00
	Lng SQ80	Squamous cell carcinoma	Lung 14	0.03	0.00
	Lng SQ16	Squamous cell carcinoma	Lung 15	0.48	0.00
	Lng SQ79	Squamous cell carcinoma	Lung 16	0.00	0.00
15	Lng 47XQ	Squamous cell carcinoma	Lung 17	0.00	0.00
	Lng SQ43	Squamous cell carcinoma	Lung 18	0.00	0.00
	Bld 46XK		Bladder 1	0.00	0.00
	Bld TR14		Bladder 2	0.00	0.54
	Blad66X		Bladder 3	0.000	0.00
20	ClnAS43		Colon 1	34.13	0.00
	Cln AS45		Colon 2	0.00	0.00
	Cln AS46		Colon 3	0.04	0.00
	Cln AS67		Colon 4	2.27	0.01
	Cln AS89		Colon 5	0.15	1.61
25	Cln DC63		Colon 6	0.02	0.00
	Endo 68X		Endometrium1	0.11	0.22
	Endo 12XA		Endometrium2	0.03	0.06
	Kid6XD		Kidney1	0.01	0.03
	Kid710K		Kidney2	0.00	0.00
30	Liv 15XA		Liver 1	0.06	0.01
	Liv201L		Liver2	0.00	0.03
	Mam986		Mammary1	0.00	0.00
	Sto 288S		Stomach1	0.00	0.00
	Sto531S		Stomach2	0.00	0.02
35	Tst39X		Testis1	0.03	0.07
	Tst 647T		Testis 2	0.02	0.09
	Thr590D		Thyroid 1	0.01	0.00
	Utr135XO		Uterus 2	0.00	1.25

0= Negative

40 In the analysis of matching samples, higher expression of lng150 is detected in lung samples showing a relatively high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal pooled samples (Table 1).

45 Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher

levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows differential expression of Lng150 in 18 lung cancer tissues compared with their respective normal adjacent tissue.

5 Altogether, the high level of tissue specificity, plus the mRNA differential expression in the lung matching samples tested are believed to make Lng150 a good marker for diagnosing, monitoring, staging, imaging and treating lung cancer.

10 **DNA sequence for Lng150**

**Primers Used for QPCR Expression Analysis**

**Forward primer**

ATGGGCAGGTCTTCTTTCC (SEQ ID NO:68)

**Reverse primer**

15 AGGCAGTTCTGTTACCCCCTA (SEQ ID NO:69)

**Probe**

TGTGCTAAGGACAGGATTGGTTGGGTA (SEQ ID NO:70)

**Example 18**

**Sequence 18**

20 **Lng141**

Gene ID 20152

ddx lung code SQLng003

**Table 1.** The absolute numbers are relative levels of expression of Lng141 in 24 normal different tissues. All 25 the values are compared to normal brain (calibrator). These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

Tissue	NORMAL
Adrenal Gland	0.04
Bladder	0.00
Brain	1.00

	Cervix	0.77
	Colon	0.02
	Endometrium	0.36
	Esophagus	0.00
5	Heart	0.02
	Kidney	0.05
	Liver	0.00
	Lung	3.45
	Mammary Gland	0.99
10	Muscle	0.31
	Ovary	2.23
	Pancreas	0.06
	Prostate	0.31
	Rectum	0.65
15	Small Intestine	0.04
	Spleen	0.70
	Stomach	0.07
	Testis	0.28
	Thymus	0.91
20	Trachea	0.69
	Uterus	1.27

0=negative

The relative levels of expression in Table 1 show that Lng141 mRNA expression is high in lung compared with most 25 other normal tissues analyzed.

The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue 30 samples of a single individual in Table 2.

**Table 2.** The absolute numbers are relative levels of expression of Lng141 in 50 pairs of matching samples. All the values are compared to normal brain (calibrator). A matching pair is formed by mRNA from the cancer sample for 35 a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

	Sample ID	Cancer Type	Tissue	CANCER	MATCHING NORMAL ADJACENT
	Lng 60L	Adenocarcinoma	Lung 1	7.14	1.89
	Lng 143L	Adenocarcinoma	Lung 2	0.11	0.18
	Lng 60XL	Adenocarcinoma	Lung 3	0.19	0.39
40	Lng AC82	Adenocarcinoma	Lung 4	0.29	0.00
	Lng AC66	Adenocarcinoma	Lung 5	5.06	1.62

5	Lng AC69	Adenocarcinoma	Lung 6	27.00	4.17
	Lng AC11	Adenocarcinoma	Lung 7	1.52	3.43
	Lng AC32	Adenocarcinoma	Lung 8	2.14	2.94
	Lng AC94	Adenocarcinoma	Lung 9	2.45	4.35
	Lng 223L	Adenocarcinoma	Lung 10	0.00	1.21
	Lng BR26	Bronchio-alveolar carcinoma	Lung 11	1.15	0.00
	Lng BA641	Bronchogenic carcinoma	Lung 12	19.84	0.80
	Lng SQ45	Squamous cell carcinoma	Lung 13	31.02	7.78
	Lng SQ14	Squamous cell carcinoma	Lung 14	0.29	0.58
	Lng SQ9X	Squamous cell carcinoma	Lung 15	0.54	0.50
10	Lng SQ56	Squamous cell carcinoma	Lung 16	1.08	0.52
	Lng SQ80	Squamous cell carcinoma	Lung 17	0.55	0.85
	Lng SQ32	Squamous cell carcinoma	Lung 18	0.68	0.91
	Lng SQ16	Squamous cell carcinoma	Lung 19	0.90	0.79
	Lng SQ79	Squamous cell carcinoma	Lung 20	8.11	6.87
15	Lng 90X	Squamous cell carcinoma	Lung 21	0.00	0.38
	Lng 47XQ	Squamous cell carcinoma	Lung 22	0.24	0.28
	Lng BR94	Squamous cell carcinoma	Lung 23	0.30	0.00
	Lng SQ43	Squamous cell carcinoma	Lung 24	2.86	0.19
	Lng LC71	Large cell carcinoma	Lung 25	0.62	2.30
20	Lng LC109	Large cell carcinoma	Lung 26	0.09	1.33
	Lng MT67	Metastasis from renal carcinoma	Lung 27	0.63	0.67
	Bld TR14		Bladder 2	0.00	0.00
25	Bld 46XK		Bladder 3	0.00	0.00
	Cln AS89		Colon 1	2.40	11.55
	Cln AS67		Colon 2	1.43	1.89
	Cln AS45		Colon 3	0.00	0.27
	Cln AS46		Colon 4	0.00	0.00
30	Cln AS43		Colon 5	6.25	0.00
	End 28XA		Endometrium1	1.25	1.59
	kid 10XD		Kidney1	0.28	1.02
	Kid 109XD		Kidney2	0.23	0.64
	Liv 15XA		Liver 1	1.22	0.84
35	Mam173M		Mammary 1	0.19	0.47
	Mam 220		Mammary 2	0.00	0.31
	Mam 355		Mammary 3	0.36	0.09
	Ovr A084		Ovary 1	3.22	0.94
	Pro 101XB		Prostate 1	0.55	58.28
40	Pro 109 XB		Prostate 2	0.11	0.21
	Pro 125XB		Prostate 3	0.24	0.26
	Sto 115S		Stomach 1	0.30	0.26
	Sto 264S		Stomach 2	0.35	0.26
	Sto 288S		Stomach 3	0.06	0.00
45	Tst 647T		Testis 1	6.87	3.96
	Utr 135XO		Uterus 1	0.00	6.02

0= Negative

In the analysis of matching samples, higher expression of Lng141 is detected in lung samples showing a relatively high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained 5 with normal pooled samples (Table 1).

Furthermore, we compared the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual. This comparison provides an indication of specificity for the cancer stage (e.g. higher 10 levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows differential expression of Lng141 in 27 lung cancer tissues compared with their respective normal adjacent tissue.

Altogether, the high level of tissue specificity, plus 15 the mRNA differential expression in the lung matching samples tested are believed to make Lng141 a good marker for diagnosing, monitoring, staging, imaging and treating lung cancer.

**DNA sequence for Lng141**

20 Sequence available from Incyte database.

**Primers Used for QPCR Expression Analysis**

**Forward primer**

ACTGCCACCACGCTTTATA (SEQ ID NO:71)

**Reverse primer**

25 TGAGGGTGGGGAGAGGTTAC (SEQ ID NO:72)

**Probe**

AGTCACATTATTAGAGGTTCGCATCTCAGG (SEQ ID NO:73)